GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 19, 2004, 02:49:51 ; Search time 74 Seconds
(without alignments)
427.462 Million cell updates/sec

Title: RIBOZYME1A
Perfect score: 57
Sequence: 1 ggguccaccuccucgcggun.....uucgcauggcuaagggaccc 57

Scoring table: IDENTITY_NUCDX
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		31	54.4	61	1	US-08-238-963A-18	Sequence 18, Appl
2		29.6	51.9	52	1	US-08-741-881-40	Sequence 40, Appl
3		29.6	51.9	52	1	US-08-741-881-80	Sequence 80, Appl
4		29.6	51.9	52	1	US-08-739-158-40	Sequence 40, Appl
5		29.6	51.9	52	1	US-08-739-158-80	Sequence 80, Appl
6		29.6	51.9	52	2	US-08-739-167-40	Sequence 40, Appl
7		29.6	51.9	52	2	US-08-739-167-80	Sequence 80, Appl
8		29.6	51.9	52	3	US-08-404-796-40	Sequence 40, Appl
9		29.6	51.9	52	3	US-08-404-796-80	Sequence 80, Appl
10		29.6	51.9	52	3	US-08-931-869-40	Sequence 40, Appl
11		29.6	51.9	52	3	US-08-931-869-80	Sequence 80, Appl
12		29.6	51.9	52	4	US-09-350-399-40	Sequence 40, Appl
13		29.6	51.9	52	4	US-09-350-399-80	Sequence 80, Appl
14		29.6	51.9	52	4	US-09-236-140A-40	Sequence 40, Appl
15		29.6	51.9	52	4	US-09-236-140A-80	Sequence 80, Appl
16		29.6	51.9	80	1	US-08-238-963A-16	Sequence 16, Appl
17		29.6	51.9	83	3	US-08-646-695-12	Sequence 12, Appl
18		29.6	51.9	83	5	PCT-US96-06053-12	Sequence 12, Appl
19		29.6	51.9	84	1	US-08-238-963A-13	Sequence 13, Appl
20		29.6	51.9	84	1	US-08-238-963A-14	Sequence 14, Appl
21		29.6	51.9	86	1	US-08-238-963A-6	Sequence 6, Appli
22		29.6	51.9	90	4	US-09-554-337-2	Sequence 2, Appli
23		29.6	51.9	90	4	US-09-733-042-12	Sequence 12, Appl
24	c	29.6	51.9	90	4	US-09-733-042-13	Sequence 13, Appl
25		29.6	51.9	101	6	5225337-2	Patent No. 5225337
26	c	29.6	51.9	105	2	US-08-553-619B-25	Sequence 25, Appl
27	c	29.6	51.9	105	2	US-08-553-619B-30	Sequence 30, Appl
28		29.6	51.9	8100	4	US-09-554-337-4	Sequence 4, Appli
29		29.6	51.9	14311	3	US-08-646-695-1	Sequence 1, Appli
30	c	29.6	51.9	14311	3	US-08-646-695-7	Sequence 7, Appli
31		29.6	51.9	14311	5	PCT-US96-06053-1	Sequence 1, Appli
32	c	29.6	51.9	14311	5	PCT-US96-06053-7	Sequence 7, Appli
33		29.6	51.9	15538	4	US-09-554-337-1	Sequence 1, Appli
34	c	29.2	51.2	601	4	US-09-814-951A-6	Sequence 6, Appli
35	c	29.2	51.2	1445	4	US-09-814-951A-1	Sequence 1, Appli
36	c	29.2	51.2	9704	4	US-09-814-951A-3	Sequence 3, Appli
37		28.8	50.5	85	1	US-08-238-963A-12	Sequence 12, Appl
38		28.4	49.8	1197	4	US-09-579-182-1	Sequence 1, Appli
39	c	28.4	49.8	1600	3	US-08-602-791-1	Sequence 1, Appli
40	c	28.4	49.8	3386	3	US-09-041-886-1	Sequence 1, Appli
41		28.2	49.5	56	4	US-09-415-784-39	Sequence 39, Appl
42		28.2	49.5	56	4	US-09-415-785A-39	Sequence 39, Appl
43		28.2	49.5	56	4	US-08-944-465-39	Sequence 39, Appl
44		28.2	49.5	56	4	US-09-415-868-39	Sequence 39, Appl
45		28.2	49.5	56	4	US-09-415-900-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-08-238-963A-18
; Sequence 18, Application US/08238963A
; Patent No. 5625047
; GENERAL INFORMATION:
; APPLICANT: Been, Michael D.
; APPLICANT: Rosenstein, Sarah P.
; APPLICANT: Perrotta, Anne T.
; TITLE OF INVENTION: ENZYMATIC RNA MOLECULES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,963A
; FILING DATE: May 5, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/821,155
; FILING DATE: January 13, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 207/093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-238-963A-18

Query Match 54.4%; Score 31; DB 1; Length 61;
Best Local Similarity 77.4%; Pred. No. 1.4;
Matches 41; Conservative 6; Mismatches 2; Indels 4; Gaps 1;

```
Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACC 56
              ||||||||||||||||::::::|||||||    |||||  |||||||||||| |
Db          8 UCCACCUCCUCGCGGUCCGACCUGGGCAU----CUUCGGAUGGCUAAGGGAGC 56
```

RESULT 2
US-08-741-881-40
; Sequence 40, Application US/08741881
; Patent No. 5789245
; GENERAL INFORMATION:
; APPLICANT: Dubensky Jr, Thomas W
; APPLICANT: Polo, John M.
; APPLICANT: Ibanez, Carlos E.
; APPLICANT: Chang, Stephen M.W.
; APPLICANT: Jolly, Douglas J.
; APPLICANT: Driver, David A.
; APPLICANT: Belli, Barbara A.
; TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,881
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 930049.423C6 / 1146.007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-741-881-40

Query Match 51.9%; Score 29.6; DB 1; Length 52;
Best Local Similarity 59.4%; Pred. No. 3.7;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32
```

RESULT 3
US-08-741-881-80
; Sequence 80, Application US/08741881
; Patent No. 5789245
; GENERAL INFORMATION:
; APPLICANT: Dubensky Jr, Thomas W
; APPLICANT: Polo, John M.
; APPLICANT: Ibanez, Carlos E.
; APPLICANT: Chang, Stephen M.W.
; APPLICANT: Jolly, Douglas J.
; APPLICANT: Driver, David A.
; APPLICANT: Belli, Barbara A.
; TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,881
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 930049.423C6 / 1146.007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-741-881-80

Query Match 51.9%; Score 29.6; DB 1; Length 52;
Best Local Similarity 59.4%; Pred. No. 3.7;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32
```

RESULT 4
US-08-739-158-40
; Sequence 40, Application US/08739158
; Patent No. 5814482
; GENERAL INFORMATION:
; APPLICANT: Dubensky Jr, Thomas W
; APPLICANT: Polo, John M.
; APPLICANT: Jolly, Douglas J.
; APPLICANT: Driver, David A.
; TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/739,158
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 930049.423D3 / 1146.012
; TELECOMMUNICATION INFORMATION:

```
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  40:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  52 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-739-158-40

  Query Match             51.9%;  Score 29.6;  DB 1;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 5
US-08-739-158-80
; Sequence 80, Application US/08739158
; Patent No. 5814482
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,158
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423D3 / 1146.012
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-739-158-80

  Query Match             51.9%;  Score 29.6;  DB 1;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 6
US-08-739-167-40
; Sequence 40, Application US/08739167
; Patent No. 5843723
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS  AND ALPHAVIRUS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,167
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C7 / 1146.008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-739-167-40

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 7
US-08-739-167-80
; Sequence 80, Application US/08739167
; Patent No. 5843723
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS  AND ALPHAVIRUS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,167
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C7 / 1146.008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-739-167-80

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 8
US-08-404-796-40
; Sequence 40, Application US/08404796
; Patent No. 6015686
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,796
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-404-796-40

  Query Match             51.9%;  Score 29.6;  DB 3;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 9
US-08-404-796-80
; Sequence 80, Application US/08404796
; Patent No. 6015686
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,796
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-404-796-80

  Query Match             51.9%;  Score 29.6;  DB 3;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 10
US-08-931-869-40
; Sequence 40, Application US/08931869
; Patent No. 6015694
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
```

```
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,869
;      FILING DATE:  16-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/404,796
;      FILING DATE:  15-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-931-869-40

  Query Match             51.9%;  Score 29.6;  DB 3;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 11
US-08-931-869-80
; Sequence 80, Application US/08931869
; Patent No. 6015694
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,869
;      FILING DATE:  16-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/404,796
;      FILING DATE:  15-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-931-869-80

  Query Match             51.9%;  Score 29.6;  DB 3;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 12
US-09-350-399-40
; Sequence 40, Application US/09350399
; Patent No. 6342372
;  GENERAL INFORMATION:
;       APPLICANT: Dubensky Jr, Thomas W
;                  Polo, John M.
;                  Jolly, Douglas J.
;                  Driver, David A.
;       TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;       NUMBER OF SEQUENCES: 128
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: US
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,399
;            FILING DATE: 08-Jul-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McMasters, David D.
;            REGISTRATION NUMBER: 33,963
;            REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;  INFORMATION FOR SEQ ID NO: 40:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 52 base pairs
;            TYPE: nucleic acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-350-399-40

  Query Match             51.9%;  Score 29.6;  DB 4;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||:::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 13
US-09-350-399-80
; Sequence 80, Application US/09350399
; Patent No. 6342372
;  GENERAL INFORMATION:
;       APPLICANT: Dubensky Jr, Thomas W
;                  Polo, John M.
;                  Jolly, Douglas J.
;                  Driver, David A.
;       TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;       NUMBER OF SEQUENCES: 128
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: US
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,399
;            FILING DATE: 08-Jul-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McMasters, David D.
;            REGISTRATION NUMBER: 33,963
;            REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 80:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 52 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-350-399-80

  Query Match             51.9%;  Score 29.6;  DB 4;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||:::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 14
US-09-236-140A-40
; Sequence 40, Application US/09236140A
; Patent No. 6376236
;  GENERAL INFORMATION:
;       APPLICANT: Dubensky Jr, Thomas W
;                  Polo, John M.
;                  Ibanez, Carlos E.
;                  Chang, Stephen M.W.
;                  Jolly, Douglas J.
;                  Driver, David A.
;                  Belli, Barbara A.
;       TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS PARTICLES
;       NUMBER OF SEQUENCES: 124
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OPPENHEIMER WOLFF & DONNELLY
;            STREET: 840 NEWPORT CENTER DRIVE, SUITE 700
;            CITY: NEWPORT BEACH
;            STATE: CALIFORNIA
;            COUNTRY: US
;            ZIP: 92660
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/236,140A
;            FILING DATE: 22-Jan-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cullman, Louis C.
;            REGISTRATION NUMBER: 39,645
;            REFERENCE/DOCKET NUMBER: 20263.332 / 1146.020
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (949) 823.6000
;            TELEFAX: (949) 823.6100
;   INFORMATION FOR SEQ ID NO: 40:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 52 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-236-140A-40

  Query Match             51.9%;  Score 29.6;  DB 4;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||:::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 15
US-09-236-140A-80
; Sequence 80, Application US/09236140A
; Patent No. 6376236
;  GENERAL INFORMATION:
;       APPLICANT: Dubensky Jr, Thomas W
;                  Polo, John M.
;                  Ibanez, Carlos E.
;                  Chang, Stephen M.W.
;                  Jolly, Douglas J.
;                  Driver, David A.
;                  Belli, Barbara A.
;       TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS PARTICLES
;       NUMBER OF SEQUENCES: 124
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OPPENHEIMER WOLFF & DONNELLY
;            STREET: 840 NEWPORT CENTER DRIVE, SUITE 700
;            CITY: NEWPORT BEACH
;            STATE: CALIFORNIA
;            COUNTRY: US
;            ZIP: 92660
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/236,140A
;             FILING DATE: 22-Jan-1999
;             CLASSIFICATION: <Unknown>
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Cullman, Louis C.
;             REGISTRATION NUMBER: 39,645
;             REFERENCE/DOCKET NUMBER: 20263.332 / 1146.020
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (949) 823.6000
;             TELEFAX: (949) 823.6100
;    INFORMATION FOR SEQ ID NO: 80:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-236-140A-80

  Query Match             51.9%;  Score 29.6;  DB 4;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 3.7;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||:::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


Search completed: May 19, 2004, 04:52:40
Job time : 88 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 19, 2004, 02:46:06 ; Search time 3211 Seconds
(without alignments)
530.098 Million cell updates/sec

Title: RIBOZYME1A
Perfect score: 57
Sequence: 1 ggguccaccuccucgcggun.....uucgcauggcuaagggaccc 57

Scoring table: IDENTITY_NUCDX
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	32.2	56.5	857	14	CF552723	CF552723 AGENCOURT
c	2	32	56.1	1182	12	BI457296	BI457296 603185617
	3	31.4	55.1	622	29	CC526419	CC526419 CH240_401
c	4	30.6	53.7	869	14	CB991305	CB991305 AGENCOURT
	5	30.2	53.0	382	10	AW681122	AW681122 WS1_8_E02
	6	30.2	53.0	405	10	AW681043	AW681043 WS1_8_E02
	7	30.2	53.0	448	10	AW283452	AW283452 LG1_272_E
	8	30.2	53.0	726	10	BF686394	BF686394 602143774
	9	30.2	53.0	807	10	BE902113	BE902113 601674890
	10	29.8	52.3	538	10	BE682984	BE682984 181389 MA
c	11	29.8	52.3	777	10	BF796052	BF796052 602259070
c	12	29.8	52.3	808	12	BM006402	BM006402 603615114
	13	29.8	52.3	823	29	CC593873	CC593873 CH240_395
c	14	29.4	51.6	495	12	BJ004361	BJ004361 BJ004361
c	15	29.4	51.6	513	10	BE386169	BE386169 601274411
	16	29.4	51.6	1057	12	BM553956	BM553956 AGENCOURT
c	17	29.2	51.2	123	9	AU077016	AU077016 AU077016
c	18	29.2	51.2	222	14	CF122359	CF122359 UI-HF-CB0
c	19	29.2	51.2	257	12	BM798569	BM798569 K-EST0082
c	20	29.2	51.2	300	9	AU099987	AU099987 AU099987
c	21	29.2	51.2	349	14	CB126416	CB126416 K-EST0175
c	22	29.2	51.2	370	9	AL701336	AL701336 DKFZp686L
c	23	29.2	51.2	380	14	CB144469	CB144469 K-EST0198
c	24	29.2	51.2	428	14	CF144607	CF144607 UI-HF-CB0
c	25	29.2	51.2	453	12	BM836423	BM836423 K-EST0112
c	26	29.2	51.2	453	14	CB126144	CB126144 K-EST0175
c	27	29.2	51.2	457	14	CF135198	CF135198 UI-HF-CB0
c	28	29.2	51.2	481	14	CB113346	CB113346 K-EST0155
c	29	29.2	51.2	487	9	AW004672	AW004672 ws91f12.x
c	30	29.2	51.2	496	14	CB125500	CB125500 K-EST0174
c	31	29.2	51.2	496	14	CB127743	CB127743 K-EST0177
c	32	29.2	51.2	501	10	BE386170	BE386170 601274413
c	33	29.2	51.2	503	10	BE269642	BE269642 601185048
c	34	29.2	51.2	507	14	CB140642	CB140642 K-EST0194
c	35	29.2	51.2	508	10	BE296622	BE296622 601173643
	36	29.2	51.2	511	12	BI053587	BI053587 MR3-GN046
c	37	29.2	51.2	513	14	CB141699	CB141699 K-EST0195
c	38	29.2	51.2	513	28	AQ294266	AQ294266 HS_3009_A
c	39	29.2	51.2	516	10	BE296669	BE296669 601174409
c	40	29.2	51.2	516	14	CB114356	CB114356 K-EST0157
c	41	29.2	51.2	519	14	CB125200	CB125200 K-EST0173
c	42	29.2	51.2	529	14	CB125108	CB125108 K-EST0173
c	43	29.2	51.2	529	14	CB146530	CB146530 K-EST0201
c	44	29.2	51.2	533	10	BF304959	BF304959 601888662
c	45	29.2	51.2	536	14	CB126208	CB126208 K-EST0175

ALIGNMENTS

RESULT 1
CF552723/c
LOCUS CF552723 857 bp mRNA linear EST 22-SEP-2003
DEFINITION AGENCOURT_15595333 NIH_MGC_183 Homo sapiens cDNA clone IMAGE:30529107 5', mRNA sequence.
ACCESSION CF552723
VERSION CF552723.1 GI:34889557
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

```
           http://image.llnl.gov
           Plate: NDAM616  row: c  column: 04
           High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .857
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529107"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match             56.5%;  Score 32.2;  DB 14;  Length 857;
  Best Local Similarity   50.9%;  Pred. No. 1e+03;
  Matches   29;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||: | ||:||:|| |  :::::::|| ||:|  ||: ||| :||  |  || |||
Db        178 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGGTGAGCGGGCCC 122


RESULT 2
BI457296/c
LOCUS       BI457296               1182 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603185617F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258380 5',
            mRNA sequence.
ACCESSION   BI457296
VERSION     BI457296.1  GI:15247965
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1182)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1870  row: a  column: 05
            High quality sequence stop: 335.
FEATURES             Location/Qualifiers
     source          1. .1182
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5258380"
                     /tissue_type="epithelioid carcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_42"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
ORIGIN
```

```
  Query Match             56.1%;  Score 32;  DB 12;  Length 1182;
  Best Local Similarity   52.3%;  Pred. No. 1.3e+03;
  Matches   23;  Conservative   14;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGC 47
              :|| | :||:|||||:::::::: || |: ||||::|||  |||
Db        827 TCCTCGTCCTCGCGGTCGGTTGTTGGGATTCGGCTTCGCTGGGC 784


RESULT 3
CC526419
LOCUS       CC526419                622 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_401C15.TARBAC13P2 CHORI-240 Bos taurus genomic clone
            CH240_401C15, genomic survey sequence.
ACCESSION   CC526419
VERSION     CC526419.1  GI:31844707
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_401C15.T7
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 401  row: C  column: 15
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .622
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_401C15"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             55.1%;  Score 31.4;  DB 29;  Length 622;
  Best Local Similarity   50.9%;  Pred. No. 1.5e+03;
  Matches   27;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 CCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              || ||:||:|||||:::::: |||| : |  |: | | :| |: |||  |||
Db        104 CCTCCTCCTCGCGGTGTGGATGGGGCGTCCATCTCCACCTGCCTCAGGCCCCC 156
```

```
RESULT 4
CB991305/c
LOCUS       CB991305                869 bp    mRNA    linear   EST 01-MAY-2003
DEFINITION  AGENCOURT_13666258 NIH_MGC_148 Homo sapiens cDNA clone
            IMAGE:30332633 5', mRNA sequence.
ACCESSION   CB991305
VERSION     CB991305.1  GI:30285837
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 869)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM349  row: p  column: 18
            High quality sequence stop: 489.
FEATURES             Location/Qualifiers
     source          1. .869
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30332633"
                     /tissue_type="pre-eclamptic placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_148"
                     /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                     alI-XhoI; Site_2: BamH; Library is oligo-dT primed and
                     directionally cloned using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert
                     size 2.3 kb  and normalized to ROT 5. This is a primary
                     library enriched for  full-lenght clones and constructed
                     using the Cap-trapper method  (Carninci, in preparation).
                     Library constructed by M. Brownstein  (NIMH/NHGRI,
                     National Institutes of Health). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             53.7%;  Score 30.6;  DB 14;  Length 869;
  Best Local Similarity   51.0%;  Pred. No. 2.8e+03;
  Matches   25;  Conservative   13;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUA 49
              |||: | ||:||:|| |  :::::::|| ||:|  ||: ||| :||| |
Db        370 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGGCGA 322


RESULT 5
AW681122
LOCUS       AW681122                382 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  WS1_8_E02.g1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW681122
VERSION     AW681122.1  GI:7554830
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
            Pratt,L.H.
  TITLE     An EST database from Sorghum: water-stressed plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence stop: 382
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Water-stressed 1 (WS1)"
                     /note="Organ: Mix of 5-week old plants on days 7 & 8 after
                     water was withheld; Vector: Lambda Zap; Site_1: XhoI;
                     Site_2: EcoRI; The library was made from poly-A RNA in the
                     cloning vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

  Query Match             53.0%;  Score 30.2;  DB 10;  Length 382;
  Best Local Similarity   49.1%;  Pred. No. 2.6e+03;
  Matches   27;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGAC 55
              || :||| |:||:|   | :::::::||     ||  :: |||:||| |||||||
Db         10 GGTTCCAGCTCCTCCACGACCTGGCTGGCTGATCGCATTAGCATGGCAAAGGGAC 64


RESULT 6
AW681043
LOCUS       AW681043                405 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  WS1_8_E02.b1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW681043
VERSION     AW681043.1  GI:7554749
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
            Pratt,L.H.
  TITLE     An EST database from Sorghum: water-stressed plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: JEN REV
            High quality sequence stop: 405
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Sorghum bicolor"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Water-stressed 1 (WS1)"
                     /note="Organ: Mix of 5-week old plants on days 7 & 8 after
                     water was withheld; Vector: Lambda Zap; Site_1: XhoI;
                     Site_2: EcoRI; The library was made from poly-A RNA in the
                     cloning vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

  Query Match             53.0%;  Score 30.2;  DB 10;  Length 405;
  Best Local Similarity   49.1%;  Pred. No. 2.6e+03;
  Matches   27;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGAC 55
              || :||| |:||:|   | :::::::||     ||  :: |||:||| |||||||
Db         10 GGTTCCAGCTCCTCCACGACCTGGCTGGCTGATCGCATTAGCATGGCAAAGGGAC 64


RESULT 7
AW283452
LOCUS       AW283452                448 bp    mRNA    linear   EST 17-DEC-2001
DEFINITION  LG1_272_E01.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW283452
VERSION     AW283452.3  GI:17886563
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     On Jan 6, 2000 this sequence version replaced gi:6857936.
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for high quality sequence is
            20. Three-prime sequences, which are obtained with PolyTMix or T7
            sequencing primer, are presented as the reverse complement.
            Seq primer: T7
            High quality sequence start: 3
            High quality sequence stop: 448
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

  Query Match             53.0%;  Score 30.2;  DB 10;  Length 448;
  Best Local Similarity   49.1%;  Pred. No. 2.8e+03;
  Matches   27;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGAC 55
              || :||| |:||:|   | :::::::||     ||  :: |||:||| |||||||
Db         75 GGTTCCAGCTCCTCCACGACCTGGCTGGCTGATCGCATTAGCATGGCAAAGGGAC 129


RESULT 8
BF686394
LOCUS       BF686394                726 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602143774F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304611 5',
            mRNA sequence.
ACCESSION   BF686394
VERSION     BF686394.1  GI:11971802
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1170  row: d  column: 20
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4304611"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             53.0%;  Score 30.2;  DB 10;  Length 726;
  Best Local Similarity   54.9%;  Pred. No. 3.4e+03;
  Matches   28;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 GGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGG 52
              | :|  || ||  || |:::: :::||||| |||||: | ||:||| | ||
Db        647 GCTCGCCCACCCTGCTGTGCTCAATGGGCAGGCGGCTGCACATGGCCAGGG 697


RESULT 9
BE902113
LOCUS       BE902113                807 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  601674890F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3957479 5',
            mRNA sequence.
ACCESSION   BE902113
VERSION     BE902113.1  GI:10391968
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM835  row: d  column: 24
            High quality sequence stop: 778.
FEATURES              Location/Qualifiers
     source           1. .807
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3957479"
                      /tissue_type="choriocarcinoma"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_21"
                      /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                      Site_2: EcoRI; cDNA made by oligo-dT priming.
                      Directionally cloned into EcoRI/XhoI sites using the
                      following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                      for average insert size 1.8kb.  Library constructed by
                      Ling Hong in the laboratory of Gerald M. Rubin (University
                      of California, Berkeley) using  ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match             53.0%;  Score 30.2;  DB 10;  Length 807;
  Best Local Similarity   54.9%;  Pred. No. 3.5e+03;
  Matches   28;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 GGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGG 52
              | :|  || ||: || | ::: :::||||| |||||: | ||:||| | ||
Db        624 GCTCGCCCACCTGGCTGCTGTCAATGGGCAGGCGGCTGCACATGGCCAGGG 674


RESULT 10
BE682984
LOCUS       BE682984                538 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  181389 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE682984
VERSION     BE682984.1  GI:10069395
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
   PUBMED   11282978
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 84  row: O  column: 20
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .538
                      /organism="Bos taurus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9913"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="MARC 4BOV"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from day 20 and day 40
                      embryos."
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 10;  Length 538;
  Best Local Similarity   46.7%;  Pred. No. 3.8e+03;
  Matches   21;  Conservative   15;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCU 48
              : ||||:||  | | :::::::|||| :|| ||::| | :|||:
Db        393 TGCACCTCCCTGAGCTTCCTGGTGGGCTTGCAGCTTCCCCTGGCT 437


RESULT 11
BF796052/c
LOCUS       BF796052                777 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602259070F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342409 5',
            mRNA sequence.
ACCESSION   BF796052
VERSION     BF796052.1  GI:12101106
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9957  row: k  column: 18
            High quality sequence stop: 716.
FEATURES              Location/Qualifiers
     source           1. .777
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4342409"
                      /tissue_type="lymphoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_85"
                      /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                      Average insert size 1.867 kb. Library enriched for
                      full-length clones and constructed by Life Technologies.
                      Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 10;  Length 777;
  Best Local Similarity   50.9%;  Pred. No. 4.5e+03;
  Matches   27;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACC 56
              :|||||:||:|  || :::::  | | :| ||| :  ||:|||  |||| ||
```

```
Db        268 TCCACCTCCTCCTGGCTGAGCACCGTCCTGGGGCCTTCCATGGCAGAGGGTCC 216


RESULT 12
BM006402/c
LOCUS       BM006402                808 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603615114F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5421075 5',
            mRNA sequence.
ACCESSION   BM006402
VERSION     BM006402.1  GI:16520756
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1876  row: l  column: 04
            High quality sequence stop: 552.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5421075"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 12;  Length 808;
  Best Local Similarity   50.9%;  Pred. No. 4.5e+03;
  Matches   27;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACC 56
              :|||||:||:|  || ::::::  | | :| ||| :  ||:|||  |||| ||
Db        256 TCCACCTCCTCCTGGCTGAGCACCGTCCTGGGGCCTTCCATGGCAGAGGGTCC 204


RESULT 13
CC593873
LOCUS       CC593873                823 bp    DNA     linear   GSS 18-JUN-2003
DEFINITION  CH240_395G6.TARBAC13P2 CHORI-240 Bos taurus genomic clone
            CH240_395G6, genomic survey sequence.
ACCESSION   CC593873
VERSION     CC593873.1  GI:31954012
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 823)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_395G6.T7
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 395  row: G  column: 6
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .823
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_395G6"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 29;  Length 823;
  Best Local Similarity   46.7%;  Pred. No. 4.6e+03;
  Matches   21;  Conservative   15;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCU 48
              : ||||:||  | | ::::::::|||| :|| ||::| | :|||:
Db        512 TGCACCTCCCTGAGCTTCCTGGTGGGCTTGCAGCTTCCCCTGGCT 556


RESULT 14
BJ004361/c
LOCUS       BJ004361                495 bp    mRNA    linear   EST 05-DEC-2001
DEFINITION  BJ004361 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA064D03 5',
            mRNA sequence.
ACCESSION   BJ004361
VERSION     BJ004361.1  GI:17355872
KEYWORDS    EST.
SOURCE      Oryzias latipes (Japanese medaka)
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
  TITLE     Medaka EST Project in Takeda's lab
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
```

```
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .495
                      /organism="Oryzias latipes"
                      /mol_type="mRNA"
                      /strain="Hd-rR"
                      /db_xref="taxon:8090"
                      /clone="MF01SSA064D03"
                      /sex="mixture of female and male"
                      /tissue_type="whole embryo"
                      /dev_stage="segmentation stage 20 - 25"
                      /clone_lib="MF01SSA cDNA"
ORIGIN

  Query Match             51.6%;  Score 29.4;  DB 12;  Length 495;
  Best Local Similarity   47.1%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative   14;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGA 54
              :||| |:||: || |::::::: | |  :|||| ::||  :||| |  |||
Db         85 TCCAGCTCCTTGCAGTCGCTGCAGTGGGTGCGGTTTCGTGTGGCGAGTGGA 35


RESULT 15
BE386169/c
LOCUS       BE386169                 513 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  601274411F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3615363 5',
            mRNA sequence.
ACCESSION   BE386169
VERSION     BE386169.1  GI:9331534
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM279  row: f  column: 04
            High quality sequence start: 15
            High quality sequence stop: 513.
FEATURES              Location/Qualifiers
     source           1. .513
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3615363"
                      /tissue_type="melanotic melanoma"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_20"
                      /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming.  Directionally
                      cloned into EcoRI/XhoI sites using the following 5'
                      adaptor: GGCACGAG(G).  Size-selected >500bp for average
                      insert size 1.8kb.  Library constructed by Ling Hong in
                      the laboratory of Gerald M. Rubin (University of
                      California, Berkeley) using  ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match             51.6%;  Score 29.4;  DB 10;  Length 513;
  Best Local Similarity   49.0%;  Pred. No. 4.9e+03;
  Matches   25;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAG 51
              |||: | ||:||:|| |  :::::::|| ||:|  ||: ||| :||   ||
Db         58 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGGAAGAG 8


Search completed: May 19, 2004, 04:51:37
Job time : 3251 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 19, 2004, 01:34:20 ; Search time 403 Seconds
(without alignments)
600.862 Million cell updates/sec

Title: RIBOZYME1A
Perfect score: 57
Sequence: 1 ggguccaccuccucgcggun.....uucgcauggcuaagggaccc 57

Scoring table: IDENTITY_NUCDX
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56.2	98.6	57	3	AAZ57636	Aaz57636 Trans-act
2	56.2	98.6	57	3	AAZ57637	Aaz57637 Trans-act
3	46.6	81.8	57	3	AAZ57638	Aaz57638 Trans-act
4	34.6	60.7	37	3	AAZ57639	Aaz57639 Nucleotid
5	31	54.4	61	2	AAQ46667	Aaq46667 Substrate
6	29.6	51.9	51	2	AAQ86173	Aaq86173 Primer HD
7	29.6	51.9	52	2	AAQ86204	Aaq86204 Sindbis/H
8	29.6	51.9	52	2	AAT31151	Aat31151 HDV rever
9	29.6	51.9	52	2	AAT30817	Aat30817 HDV riboz
10	29.6	51.9	52	2	AAV42394	Aav42394 Nested PC
11	29.6	51.9	52	2	AAV42426	Aav42426 Forward P
12	29.6	51.9	52	2	AAV60154	Aav60154 Nested PC
13	29.6	51.9	52	2	AAV60185	Aav60185 Reverse P
14	29.6	51.9	52	2	AAV70746	Aav70746 Reverse P
15	29.6	51.9	52	2	AAV70714	Aav70714 Nested pr
16	29.6	51.9	52	3	AAZ92930	Aaz92930 Hepatitis
17	29.6	51.9	52	3	AAZ92970	Aaz92970 Hepatitis
18	29.6	51.9	52	3	AAZ92843	Aaz92843 Hepatitis
19	29.6	51.9	52	3	AAZ92803	Aaz92803 Hepatitis
20	29.6	51.9	52	6	AAL38850	Aal38850 Alphaviru
21	29.6	51.9	52	6	AAL38810	Aal38810 Alphaviru
22	29.6	51.9	52	6	ABK46276	Abk46276 HDV antig
23	29.6	51.9	52	6	ABK46316	Abk46316 Hepatitis
24	29.6	51.9	80	2	AAQ53138	Aaq53138 Substrate
25	29.6	51.9	80	2	AAQ53140	Aaq53140 Substrate
26	29.6	51.9	83	5	AAF28121	Aaf28121 HDV riboz
27	29.6	51.9	85	2	AAQ46665	Aaq46665 Self-clea
28	29.6	51.9	86	2	AAQ46663	Aaq46663 Self-clea
29	29.6	51.9	90	2	AAX76586	Aax76586 Hepatitis
30	29.6	51.9	90	5	AAF84035	Aaf84035 EPO gene
c 31	29.6	51.9	90	5	AAF84036	Aaf84036 EPO gene
32	29.6	51.9	91	7	ABZ82120	Abz82120 Hepatitis
c 33	29.6	51.9	99	7	ABZ82121	Abz82121 Hepatitis
34	29.6	51.9	567	1	AAN82174	Aan82174 Clone del
c 35	29.6	51.9	1679	1	AAN82172	Aan82172 cDNA corr
c 36	29.6	51.9	1688	7	ABZ77701	Abz77701 Nucleotid
37	29.6	51.9	5451	8	ADA41729	Ada41729 Plasmid p
38	29.6	51.9	8085	9	ADE48122	Ade48122 DNA HCV r
39	29.6	51.9	8100	2	AAX76588	Aax76588 SFV EcoRV
40	29.6	51.9	8732	9	ADE48121	Ade48121 RNA HCV r
41	29.6	51.9	12474	2	AAX78131	Aax78131 Plasmid p
42	29.6	51.9	14311	2	AAT38744	Aat38744 Plasmid p
c 43	29.6	51.9	14311	5	AAF28116	Aaf28116 Vesicular
44	29.6	51.9	14311	5	AAF28115	Aaf28115 Vesicular
45	29.6	51.9	15538	2	AAX76582	Aax76582 Plasmid p

ALIGNMENTS

RESULT 1
AAZ57636
ID AAZ57636 standard; RNA; 57 BP.
XX
AC AAZ57636;
XX
DT 05-APR-2000 (first entry)
XX
DE Trans-acting antigenomic delta ribozyme, deltaRzP1.1 nucleotide sequence.
XX
KW Trans-acting antigenomic delta ribozyme; viral RNA cleavage;
KW inherited disease; deltaRzP1.1; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_binding 1. .6
FT /*tag= a
FT /note= "Forms double stranded region with bases 52-57"
FT stem_loop 7. .19
FT /*tag= b
FT misc_binding 20. .25
FT /*tag= c
FT /bound_moiety= "SP1.1 substrate (AAZ57641) bases 5-11"
FT stem_loop 30. .45
FT /*tag= d
FT misc_binding 52. .57
FT /*tag= e
FT /note= "Forms a double stranded region with bases 6-1"
XX
PN WO9955856-A2.
XX
PD 04-NOV-1999.
XX
PF 29-APR-1999; 99WO-CA000391.
XX
PR 29-APR-1998; 98CA-02230203.
XX
PA (UYSH) UNIV SHERBROOKE.
XX
PI Perreault J, Ananvoranich S, Lafontaine D;
XX
DR WPI; 2000-096791/08.
XX
PT New construction of nucleic acid enzyme useful for biotechnological,
PT diagnostic and therapeutic applications.

```
XX
PS   Example 1; Fig 1A; 52pp; English.
XX
CC   This is the nucleotide sequence of ribozyme deltaRzP1.1. This ribozyme
CC   cleaves substrate SP1.1 (see AAZ57641). The invention relates to a
CC   nucleic acid enzyme (e.g. deltaRzP1.1) that is constructed to have a
CC   substrate binding portion with the following sequence 3'-UNNXNN-5'. The
CC   substrate of the enzyme has the sequence 5'-H'*GNNHNN-3'. The binding
CC   portion of the enzyme (ribozyme) base pairs to 6 nucleotides 3' to the
CC   cleavage site of the substrate (cleavage site is represented by *). At
CC   least one nucleotide is present 5' to the cleavage site of the substrate
CC   sequence. The enzyme of the invention is used to cleave a substrate
CC   nucleotide sequence at a specific cleavage site by mixing the substrate
CC   with the enzyme. The enzyme is used to cleave viral RNA or RNA causing
CC   for example an inherited disease. The enzymes also have other
CC   therapeutic, biotechnological and diagnostic applications
XX
SQ   Sequence 57 BP; 7 A; 21 C; 19 G; 0 T; 10 U; 0 Other;

  Query Match             98.6%;  Score 56.2;  DB 3;  Length 57;
  Best Local Similarity   89.5%;  Pred. No. 1.4e-07;
  Matches   51;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              ||||||||||||||||||::::::|||||||||||||||||||||||||||||||||
Db          1 GGGUCCACCUCCUCGCGGUCCGACCUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57


RESULT 2
AAZ57637
ID   AAZ57637 standard; RNA; 57 BP.
XX
AC   AAZ57637;
XX
DT   05-APR-2000  (first entry)
XX
DE   Trans-acting antigenomic delta ribozyme, deltaRzP1.2 nucleotide sequence.
XX
KW   Trans-acting antigenomic delta ribozyme; viral RNA cleavage;
KW   inherited disease; deltaRzP1.2; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    1. .6
FT                   /*tag= a
FT                   /note= "Forms double stranded region with bases 52-57"
FT   stem_loop       7. .19
FT                   /*tag= b
FT   misc_binding    20. .25
FT                   /*tag= c
FT                   /bound_moiety= "SP1.2 substrate bases (AAZ57634) 6-11"
FT   stem_loop       30. .45
FT                   /*tag= d
FT   misc_binding    52. .57
FT                   /*tag= e
FT                   /note= "Forms a double stranded region with bases 6-1"
XX
PN   WO9955856-A2.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-CA000391.
XX
PR   29-APR-1998;   98CA-02230203.
XX
PA   (UYSH ) UNIV SHERBROOKE.
XX
PI   Perreault J,  Ananvoranich S,  Lafontaine D;
XX
DR   WPI; 2000-096791/08.
XX
PT   New construction of nucleic acid enzyme useful for biotechnological,
PT   diagnostic and therapeutic applications.
XX
PS   Example 1; Page; 52pp; English.
XX
CC   This is the nucleotide sequence of ribozyme deltaRzP1.2. This ribozyme
CC   cleaves substrate SP1.2 (see AAZ57634). The invention relates to a
CC   nucleic acid enzyme (e.g. deltaRzP1.2) that is constructed to have a
CC   substrate binding portion with the following sequence 3'-UNNXNN-5'. The
CC   substrate of the enzyme has the sequence 5'-H'*GNNHNN-3'. The binding
CC   portion of the enzyme (ribozyme) base pairs to 6 nucleotides 3' to the
CC   cleavage site of the substrate (cleavage site is represented by *). At
CC   least one nucleotide is present 5' to the cleavage site of the substrate
CC   sequence. The enzyme of the invention is used to cleave a substrate
CC   nucleotide sequence at a specific cleavage site by mixing the substrate
CC   with the enzyme. The enzyme is used to cleave viral RNA or RNA causing
CC   for example an inherited disease. The enzymes also have other
CC   therapeutic, biotechnological and diagnostic applications. Note: This
CC   sequence is not shown in the specification, but has been derived from the
CC   deltaRzP1.1 sequence (AAZ57636) shown in figure 1
XX
SQ   Sequence 57 BP; 7 A; 21 C; 19 G; 0 T; 10 U; 0 Other;

  Query Match             98.6%;  Score 56.2;  DB 3;  Length 57;
  Best Local Similarity   89.5%;  Pred. No. 1.4e-07;
  Matches   51;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              ||||||||||||||||||::::::|||||||||||||||||||||||||||||||||
Db          1 GGGUCCACCUCCUCGCGGUCCCAGCUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57


RESULT 3
AAZ57638
ID   AAZ57638 standard; RNA; 57 BP.
XX
AC   AAZ57638;
XX
DT   15-SEP-2003  (revised)
DT   05-APR-2000  (first entry)
XX
DE   Trans-acting antigenomic delta ribozyme nucleotide sequence.
XX
KW   Trans-acting antigenomic delta ribozyme; viral RNA cleavage;
KW   inherited disease; ss.
XX
OS   Hepatitis D virus.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    1. .6
FT                   /*tag= a
FT                   /note= "Forms double stranded region with bases 52-57"
FT   stem_loop       7. .19
FT                   /*tag= b
FT   misc_binding    20. .25
FT                   /*tag= c
FT                   /bound_moiety= "SP1.1 substrate bases (AAZ57641) 5-11"
FT   stem_loop       30. .45
FT                   /*tag= d
FT   misc_binding    52. .57
FT                   /*tag= e
FT                   /note= "Forms a double stranded region with bases 6-1"
XX
PN   WO9955856-A2.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-CA000391.
XX
PR   29-APR-1998;   98CA-02230203.
XX
PA   (UYSH ) UNIV SHERBROOKE.
XX
PI   Perreault J,  Ananvoranich S,  Lafontaine D;
```

```
XX
DR   WPI; 2000-096791/08.
XX
PT   New construction of nucleic acid enzyme useful for biotechnological,
PT   diagnostic and therapeutic applications.
XX
PS   Example 1; Fig 2; 52pp; English.
XX
CC   This is the nucleotide sequence of a trans-acting antigenomic delta
CC   ribozyme of the invention. This ribozyme cleaves substrate SP1.1 (see
CC   AAZ57641). The invention relates to a nucleic acid enzyme (e.g.
CC   deltaRzP1.1) that is constructed to have a substrate binding portion with
CC   the following sequence 3'-UNNXNN-5'. The substrate of the enzyme has the
CC   sequence 5'-H'*GNNHNN-3'. The binding portion of the enzyme (ribozyme)
CC   base pairs to 6 nucleotides 3' to the cleavage site of the substrate
CC   (cleavage site is represented by *). At least one nucleotide is present
CC   5' to the cleavage site of the substrate sequence. The enzyme of the
CC   invention is used to cleave a substrate nucleotide sequence at a specific
CC   cleavage site by mixing the substrate with the enzyme. The enzyme is used
CC   to cleave viral RNA or RNA causing for example an inherited disease. The
CC   enzymes also have other therapeutic, biotechnological and diagnostic
CC   applications. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 57 BP; 7 A; 21 C; 19 G; 0 T; 10 U; 0 Other;

  Query Match             81.8%;  Score 46.6;  DB 3;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 0.00016;
  Matches   45;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||||||||||||||||||||::::::|||||||||      ||||||||||||||||||
Db          1 GGGUCCACCUCCUCGCGGUCCGACCUGGGCAUGCCUUCGGGCAUGGCUAAGGGACCC 57


RESULT 4
AAZ57639
ID   AAZ57639 standard; RNA; 37 BP.
XX
AC   AAZ57639;
XX
DT   05-APR-2000  (first entry)
XX
DE   Nucleotide sequence of bimolecular ribozyme RzA fragment.
XX
KW   Bimolecular ribozyme; viral RNA cleavage; RzA fragment;
KW   inherited disease; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    1. .6
FT                   /*tag=  a
FT                   /bound_moiety= "RzB fragment of bimolecular ribozyme"
FT                   /note= "Forms double-stranded region with bases 20-16 of
FT                   sequence AAZ57640"
FT   stem_loop       7. .19
FT                   /*tag=  b
FT   misc_binding    20. .25
FT                   /*tag=  c
FT                   /bound_moiety= "Ribozyme substrate P1.1"
FT                   /note= "Forms double-stranded region with bases 11-6 of
FT                   sequence AAZ57641"
FT   misc_binding    30. .37
FT                   /*tag=  d
FT                   /bound_moiety= "RzB fragment of bimolecular ribozyme"
FT                   /note= "Forms double-stranded region with bases 8-1 of
FT                   sequence AAZ57640"
XX
PN   WO9955856-A2.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-CA000391.
XX
PR   29-APR-1998;   98CA-02230203.
XX
PA   (UYSH ) UNIV SHERBROOKE.
XX
PI   Perreault J,  Ananvoranich S,  Lafontaine D;
XX
DR   WPI; 2000-096791/08.
XX
PT   New construction of nucleic acid enzyme useful for biotechnological,
PT   diagnostic and therapeutic applications.
XX
PS   Example 5; Fig 4; 52pp; English.
XX
CC   This is the nucleotide sequence bimolecular ribozyme RzA fragment. This
CC   ribozyme cleaves substrate SP1.1 (see AAZ57641). The invention relates to
CC   a nucleic acid enzyme (e.g. deltaRzP1.1) that is constructed to have a
CC   substrate binding portion with the following sequence 3'-UNNXNN-5'. The
CC   substrate of the enzyme has the sequence 5'-H'*GNNHNN-3'. The binding
CC   portion of the enzyme (ribozyme) base pairs to 6 nucleotides 3' to the
CC   cleavage site of the substrate (cleavage site is represented by *). At
CC   least one nucleotide is present 5' to the cleavage site of the substrate
CC   sequence. The enzyme of the invention is used to cleave a substrate
CC   nucleotide sequence at a specific cleavage site by mixing the substrate
CC   with the enzyme. The enzyme is used to cleave viral RNA or RNA causing
CC   for example an inherited disease. The enzymes also have other
CC   therapeutic, biotechnological and diagnostic applications
XX
SQ   Sequence 37 BP; 3 A; 16 C; 12 G; 0 T; 6 U; 0 Other;

  Query Match             60.7%;  Score 34.6;  DB 3;  Length 37;
  Best Local Similarity   81.1%;  Pred. No. 1;
  Matches   30;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGC 37
              |||||||||||||||||||||::::::||||||||| |
Db          1 GGGUCCACCUCCUCGCGGUCCGACCUGGGCAUGCGCC 37


RESULT 5
AAQ46667
ID   AAQ46667 standard; RNA; 61 BP.
XX
AC   AAQ46667;
XX
DT   25-MAR-2003  (revised)
DT   13-JAN-1994  (first entry)
XX
DE   Substrate-cleaving sequence ADC3.
XX
KW   Substrate; self-cleaving sequence; HDV; hepatitis delta virus; enzyme;
KW   duplex; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       7. .55
FT                   /*tag=  a
FT   misc_structure  7. .10
FT                   /*tag=  b
FT                   /label= stem_II
FT                   /note= "forms duplex with region 52. .55"
FT   stem_loop       11. .23
FT                   /*tag=  d
FT   misc_structure  11. .13
FT                   /*tag=  e
FT                   /label= stem_III
FT                   /note= "forms duplex with region 21. .23"
FT   misc_feature    21. .23
FT                   /*tag=  f
FT                   /label= stem_III
FT                   /note= "forms duplex with region 11. .13"
FT   misc_structure  24. .30
```

```
FT                   /*tag= g
FT                   /note= "RNA substrate binding region"
FT   misc_structure  33. .37
FT                   /*tag= i
FT                   /label= stem_IV
FT                   /note= "forms duplex with region 42. .45"
FT   stem_loop       34. .45
FT                   /*tag= h
FT   misc_feature    42. .45
FT                   /*tag= j
FT                   /label= stem_IV
FT                   /note= "forms duplex with region 33. .37"
FT   misc_feature    52. .55
FT                   /*tag= c
FT                   /label= stem_II
FT                   /note= "forms duplex with region 7. .10"
XX
PN   WO9314218-A1.
XX
PD   22-JUL-1993.
XX
PF   12-JAN-1993;   93WO-US000292.
XX
PR   13-JAN-1992;   92US-00821155.
XX
PA   (UYDU-) UNIV DUKE.
XX
PI   Been MD,  Rosenstein SP,  Perrota AT;
XX
DR   WPI; 1993-243233/30.
XX
PT   Nucleic acid molecule having RNA substrate-cleaving enzymatic activity -
PT   useful for cleaving specific target molecules in=vitro.
XX
PS   Disclosure; Fig 8; 55pp; English.
XX
CC   The self-cleaving sequences from genomic and antigenomic HDV can be used
CC   to develop enzymatic RNA mols. with similar properties. ADC3 is a smaller
CC   version of ADC1 (AAQ53138), wherein stem IV is shortened. Such a smaller
CC   enzymatic RNA has simplified synthesis and the potential for higher
CC   specific activity due to a higher probability that a small RNA will fold
CC   into an enzymatically active structure. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 61 BP; 10 A; 20 C; 19 G; 0 T; 12 U; 0 Other;

  Query Match             54.4%;  Score 31;  DB 2;  Length 61;
  Best Local Similarity   77.4%;  Pred. No. 16;
  Matches   41;  Conservative    6;  Mismatches    2;  Indels    4;  Gaps    1;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACC 56
              |||||||||||||||||:::::|||||||    |||||  |||||||||||| |
Db          8 UCCACCUCCUCGCGGUCCGACCUGGGCAU----CUUCGGAUGGCUAAGGGAGC 56


RESULT 6
AAQ86173
ID   AAQ86173 standard; DNA; 51 BP.
XX
AC   AAQ86173;
XX
DT   25-MAR-2003  (revised)
DT   21-NOV-1995  (first entry)
XX
DE   Primer HDV17-68 amplifies vector DNA to initiate Sindbis infection.
XX
KW   Eukaryotic layered vector initiation system; Sindbis; alphavirus; PCR;
KW   transcription initiation; non-structural protein; subgenomic fragment;
KW   RNA polymerase recognition sequence; ELVIS; gene therapy; amplify;
KW   primer; polymerase chain reaction; cystic fibrosis; ss.
XX
OS   Synthetic.
XX
PN   WO9507994-A2.
XX
PD   23-MAR-1995.
XX
PF   15-SEP-1994;   94WO-US010469.
XX
PR   15-SEP-1993;   93US-00122791.
PR   18-FEB-1994;   94US-00198450.
XX
PA   (VIAG-) VIAGENE INC.
XX
PI   Dubensky TW,  Ibanez CE,  Chang SM,  Jolly DJ,  Driver DA,  Polo JM;
XX
DR   WPI; 1995-131362/17.
XX
PT   New alpha virus vectors for gene therapy - of viral infection, cancer,
PT   auto:immune disease, etc., and as vaccines.
XX
PS   Example 2; Page 68; 260pp; English.
XX
CC   The sequences given in AAQ86167-81 are primers which were used in the
CC   generation of plasmid DNA which initiates Sindbis infection. The
CC   amplified DNA sequences were used in the construction of a eukaryotic
CC   layered vector initiation system (ELVIS) derived from Sindbis. ELVIS's
CC   comprise a 5' sequence capable of initiating transcription of an
CC   alphavirus, a nucleotide sequence encoding alphavirus non-structural
CC   proteins, a viral junction region which has been inactivated such that
CC   viral transcription of the subgenomic fragment is prevented, and an
CC   alphavirus RNA polymerase recognition sequence. Inactivation of the viral
CC   junction region prevents transcription of the subgenomic fragment making
CC   vectors such as this suitable for a wide variety of applications, eg.
CC   gene therapy for the treatment of cystic fibrosis. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 51 BP; 9 A; 20 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 51;
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||:::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 7
AAQ86204
ID   AAQ86204 standard; DNA; 52 BP.
XX
AC   AAQ86204;
XX
DT   25-MAR-2003  (revised)
DT   23-NOV-1995  (first entry)
XX
DE   Sindbis/HDV ribozyme sequence reverse primer, HDV17-68.
XX
KW   Eukaryotic layered vector initiation system; Sindbis; alphavirus; PCR;
KW   transcription initiation; non-structural protein; subgenomic fragment;
KW   RNA polymerase recognition sequence; ELVIS; gene therapy; amplify;
KW   primer; polymerase chain reaction; cystic fibrosis; ss.
XX
OS   Synthetic.
XX
PN   WO9507994-A2.
XX
PD   23-MAR-1995.
XX
PF   15-SEP-1994;   94WO-US010469.
XX
PR   15-SEP-1993;   93US-00122791.
PR   18-FEB-1994;   94US-00198450.
XX
PA   (VIAG-) VIAGENE INC.
```

```
XX
PI   Dubensky TW,  Ibanez CE,  Chang SM,  Jolly DJ,  Driver DA,  Polo JM;
XX
DR   WPI; 1995-131362/17.
XX
PT   New alpha virus vectors for gene therapy - of viral infection, cancer,
PT   auto:immune disease, etc., and as vaccines.
XX
PS   Example 7; Page 114; 260pp; English.
XX
CC   The sequences given in AAQ86200-15 are primers which were used in the
CC   production of alphavirus vectors expressing multiple heterologous genes.
CC   These vectors are eukaryotic layered vector initiation systems (ELVIS)
CC   derived from Sindbis. ELVIS's comprise a 5' sequence capable of
CC   initiating transcription of an alphavirus, a nucleotide sequence encoding
CC   alphavirus non-structural proteins, a viral junction region which has
CC   been inactivated such that viral transcription of the subgenomic fragment
CC   is prevented, and an alphavirus RNA polymerase recognition sequence.
CC   Inactivation of the viral junction region prevents transcription of the
CC   subgenomic fragment making vectors such as this suitable for a wide
CC   variety of applications, eg. gene therapy for the treatment of cystic
CC   fibrosis. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 8
AAT31151
ID   AAT31151 standard; DNA; 52 BP.
XX
AC   AAT31151;
XX
DT   12-SEP-1996  (first entry)
XX
DE   HDV reverse primer HDV17-68.
XX
KW   Alphavirus; Sindbis virus; vector; gene therapy; vaccine;
KW   polymerase chain reaction; PCR; primer; hepatitis delta virus; HDV;
KW   ribozyme; ss.
XX
OS   Synthetic.
XX
PN   WO9617072-A2.
XX
PD   06-JUN-1996.
XX
PF   30-NOV-1995;   95WO-US015490.
XX
PR   30-NOV-1994;   94US-00348472.
PR   18-JAN-1995;   95US-00376184.
PR   15-MAR-1995;   95US-00405827.
XX
PA   (CHIR ) CHIRON VIAGENE INC.
XX
PI   Dubensky TW,  Polo JM,  Ibanez CE,  Chang SMW,  Jolly DJ,  Driver DA;
PI   Belli BA;
XX
DR   WPI; 1996-277785/28.
XX
PT   New recombinant alpha-virus vectors - used to develop prods and methods
PT   for use in gene therapy and in the prodn. of vaccines.
XX
PS   Example 7; Page 120; 256pp; English.
XX
CC   Primer HDV17-68 (AAT31151) is based on hepatitis delta virus (HDV)
CC   nucleotides 839-887. It was used with reverse primer HDV49-XC (AAT30850)
CC   for the PCR amplification of HDV sequences. A second primer pair, SIN-HDV
CC   (AAT31152) and SIN276-SPE (AAT31153), is used to amplify a fragment
CC   contg. Sindbis 5'-end sequences fused to HDV sequences. Overlapping
CC   synthesis is used in a second round of PCR using primers HDV49-XC and
CC   SIN276-SPE. The resulting construct contains the expression cassette
CC   elements HDV ribozyme/Sindbis 5'-end 299 nts/Sindbis junction
CC   region/Sindbis structural protein genes/Sindbis 3'-end untranslated
CC   region, designated pd5'26s. This was inserted into pcDNA3. Packaging cell
CC   line cassettes were constructed that allow inducible expression of
CC   structural proteins via alphavirus vectors
XX
SQ   Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 9
AAT30817
ID   AAT30817 standard; DNA; 52 BP.
XX
AC   AAT30817;
XX
DT   12-SEP-1996  (first entry)
XX
DE   HDV ribozyme primer HDV17-68.
XX
KW   Alphavirus; Sindbis virus; vector; gene therapy; vaccine; primer;
KW   polymerase chain reaction; PCR; ELVIS; ribozyme; HDV;
KW   hepatitis delta virus; ss.
XX
OS   Synthetic.
XX
PN   WO9617072-A2.
XX
PD   06-JUN-1996.
XX
PF   30-NOV-1995;   95WO-US015490.
XX
PR   30-NOV-1994;   94US-00348472.
PR   18-JAN-1995;   95US-00376184.
PR   15-MAR-1995;   95US-00405827.
XX
PA   (CHIR ) CHIRON VIAGENE INC.
XX
PI   Dubensky TW,  Polo JM,  Ibanez CE,  Chang SMW,  Jolly DJ,  Driver DA;
PI   Belli BA;
XX
DR   WPI; 1996-277785/28.
XX
PT   New recombinant alpha-virus vectors - used to develop prods and methods
PT   for use in gene therapy and in the prodn. of vaccines.
XX
PS   Example 3; Page 85; 256pp; English.
XX
CC   Primer SHDV1F (AAT30816) contains a buffer sequence allowing enzyme
CC   digestion, a SacI site, and a hepatitis delta virus (HDV) ribozyme
CC   sequence. It was used with nested primer HDV17-68 (AAT30817) and reverse
CC   primer SHDV84R (AAT30818) to generate an HDV ribozyme. This antigenomic
CC   ribozyme can be placed between the polyA tract at the 3' end of a Sindbis
CC   virus-based eukaryotic layered vector initiation system, ELVIS, and the
CC   transcription terminal signals to produce alphavirus expression vectors
CC   useful in methods of gene therapy and for vaccine prodn
XX
SQ   Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
```

```
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 10
AAV42394
ID   AAV42394 standard; DNA; 52 BP.
XX
AC   AAV42394;
XX
DT   02-OCT-1998  (first entry)
XX
DE   Nested PCR primer HDV17-68 of the invention.
XX
KW   DNA alphavirus; structural protein expression; inhibit; pathogen;
KW   immune response; stimulate; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   US5789245-A.
XX
PD   04-AUG-1998.
XX
PF   30-OCT-1996;   96US-00741881.
XX
PR   15-SEP-1993;   93US-00122791.
PR   18-FEB-1994;   94US-00198450.
PR   30-NOV-1994;   94US-00348472.
PR   20-JAN-1995;   95US-00376184.
PR   15-MAR-1995;   95US-00404796.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Chang SMW,  Ibanez CE,  Jolly DJ,  Dubensky TW,  Driver DA,  Polo JM;
XX
DR   WPI; 1998-446089/38.
XX
PT   DNA alpha:virus structural protein expression cassettes - for producing
PT   recombinant alpha:virus particles.
XX
PS   Example 3; Col 69; 140pp; English.
XX
CC   PCR primers AAV42367-420 and AAV42422-54 are used in the course of the
CC   invention. The specification describes a DNA alphavirus structural
CC   protein expression cassette which comprises an inducible promoter and an
CC   alphavirus structural protein gene, where the promoter directs the
CC   expression of the alphavirus structural protein gene upon induction of
CC   the promoter within a cell, and where prior to induction within the cell,
CC   the expression cassette does not express sufficient quantities of
CC   structural proteins to be cytotoxic to a BHK cell containing the
CC   expression cassette. The products may be used to inhibit pathogens and
CC   stimulate an immune response
XX
SQ   Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 11
AAV42426
ID   AAV42426 standard; DNA; 52 BP.
XX
AC   AAV42426;
XX
DT   02-OCT-1998  (first entry)
XX
DE   Forward PCR primer HDV17-68.
XX
KW   DNA alphavirus; structural protein expression; inhibit; pathogen;
KW   immune response; stimulate; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   US5789245-A.
XX
PD   04-AUG-1998.
XX
PF   30-OCT-1996;   96US-00741881.
XX
PR   15-SEP-1993;   93US-00122791.
PR   18-FEB-1994;   94US-00198450.
PR   30-NOV-1994;   94US-00348472.
PR   20-JAN-1995;   95US-00376184.
PR   15-MAR-1995;   95US-00404796.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Chang SMW,  Ibanez CE,  Jolly DJ,  Dubensky TW,  Driver DA,  Polo JM;
XX
DR   WPI; 1998-446089/38.
XX
PT   DNA alpha:virus structural protein expression cassettes - for producing
PT   recombinant alpha:virus particles.
XX
PS   Example 7; Col 103; 140pp; English.
XX
CC   PCR primers AAV42367-420 and AAV42422-54 are used in the course of the
CC   invention. The specification describes a DNA alphavirus structural
CC   protein expression cassette which comprises an inducible promoter and an
CC   alphavirus structural protein gene, where the promoter directs the
CC   expression of the alphavirus structural protein gene upon induction of
CC   the promoter within a cell, and where prior to induction within the cell,
CC   the expression cassette does not express sufficient quantities of
CC   structural proteins to be cytotoxic to a BHK cell containing the
CC   expression cassette. The products may be used to inhibit pathogens and
CC   stimulate an immune response
XX
SQ   Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

  Query Match             51.9%;  Score 29.6;  DB 2;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 43;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 12
AAV60154
ID   AAV60154 standard; DNA; 52 BP.
XX
AC   AAV60154;
XX
DT   25-MAR-2003  (revised)
DT   04-DEC-1998  (first entry)
XX
DE   Nested PCR primer HDV17-68.
XX
KW   Eukaryotic layered vector initiation system; stimulate; immune response;
KW   Sindbis; PCR primer; ss.
XX
OS   Synthetic.
OS   Hepatitis D virus.
XX
PN   US5814482-A.
```

XX
PD 29-SEP-1998.
XX
PF 30-OCT-1996; 96US-00739158.
XX
PR 15-SEP-1993; 93US-00122791.
PR 18-FEB-1994; 94US-00198450.
PR 14-SEP-1994; 94WO-US010469.
PR 30-NOV-1994; 94US-00348472.
PR 18-JAN-1995; 95US-00376184.
PR 15-MAR-1995; 95US-00404796.
XX
PA (POLO/) POLO J M.
PA (DUBE/) DUBENSKY T W.
PA (JOLL/) JOLLY D J.
PA (DRIV/) DRIVER D A.
XX
PI Driver DA, Polo JM, Jolly DJ, Dubensky TW;
XX
DR WPI; 1998-541753/46.
XX
PT Eukaryotic layered vector initiation system - containing eukaryotic
PT promoter and heterologous antigen coding sequence, useful for stimulating
PT immune response.
XX
PS Example 3; Col 71-72; 144pp; English.
XX
CC PCR primers AAV60153-55 are used to amplify Hepatitis delta virus
CC ribozyme sequence. The product is used in the course of the invention.
CC The specifiaction describes an eukaryotic layered vector initiation
CC system, based on Sindbis. The eukaryotic layered vector initiation system
CC comprises a eukaryotic promoter 5' of viral cDNA which initiates, in a
CC susceptible target cell, 5' to 3' synthesis of RNA from the viral cDNA.
CC The RNA comprises a vector construct which autonomously amplifies in the
CC cell and expresses a heterologous nucleic acid sequence which encodes an
CC antigen or modified form that stimulates an immune response within an
CC animal. The system is useful for stimulating an immune response to an
CC antigen by introducing the vector into target cells, preferably by
CC infection in vivo, especially where the immune response is a cell
CC mediated, HLA class I-restricted or an HLA class II-restricted immune
CC response. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 51.9%; Score 29.6; DB 2; Length 52;
Best Local Similarity 59.4%; Pred. No. 43;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

Qy 4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
:|||||:||:|||||::::::::|||||: ||
Db 1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32

RESULT 13
AAV60185
ID AAV60185 standard; DNA; 52 BP.
XX
AC AAV60185;
XX
DT 25-MAR-2003 (revised)
DT 04-DEC-1998 (first entry)
XX
DE Reverse PCR primer HDV17-68.
XX
KW Eukaryotic layered vector initiation system; stimulate; immune response;
KW Sindbis; PCR primer; ss.
XX
OS Synthetic.
XX
PN US5814482-A.
XX
PD 29-SEP-1998.
XX
PF 30-OCT-1996; 96US-00739158.
XX
PR 15-SEP-1993; 93US-00122791.
PR 18-FEB-1994; 94US-00198450.
PR 14-SEP-1994; 94WO-US010469.
PR 30-NOV-1994; 94US-00348472.
PR 18-JAN-1995; 95US-00376184.
PR 15-MAR-1995; 95US-00404796.
XX
PA (POLO/) POLO J M.
PA (DUBE/) DUBENSKY T W.
PA (JOLL/) JOLLY D J.
PA (DRIV/) DRIVER D A.
XX
PI Driver DA, Polo JM, Jolly DJ, Dubensky TW;
XX
DR WPI; 1998-541753/46.
XX
PT Eukaryotic layered vector initiation system - containing eukaryotic
PT promoter and heterologous antigen coding sequence, useful for stimulating
PT immune response.
XX
PS Example 5; Col 108; 144pp; English.
XX
CC PCR primers AAV60184-85 are used in the course of the invention. The
CC specifiaction describes an eukaryotic layered vector initiation system,
CC based on Sindbis. The eukaryotic layered vector initiation system
CC comprises a eukaryotic promoter 5' of viral cDNA which initiates, in a
CC susceptible target cell, 5' to 3' synthesis of RNA from the viral cDNA.
CC The RNA comprises a vector construct which autonomously amplifies in the
CC cell and expresses a heterologous nucleic acid sequence which encodes an
CC antigen or modified form that stimulates an immune response within an
CC animal. The system is useful for stimulating an immune response to an
CC antigen by introducing the vector into target cells, preferably by
CC infection in vivo, especially where the immune response is a cell
CC mediated, HLA class I-restricted or an HLA class II-restricted immune
CC response. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 51.9%; Score 29.6; DB 2; Length 52;
Best Local Similarity 59.4%; Pred. No. 43;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

Qy 4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
:|||||:||:|||||::::::::|||||: ||
Db 1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32

RESULT 14
AAV70746
ID AAV70746 standard; DNA; 52 BP.
XX
AC AAV70746;
XX
DT 20-MAR-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Reverse PCR primer HDV17-68 used to amplify Hepatitis delta virus.
XX
KW Alphavirus vector construct; gene therapy; PCR primer; ss.
XX
OS Synthetic.
OS Hepatitis D virus.
XX
PN US5843723-A.
XX
PD 01-DEC-1998.
XX
PF 30-OCT-1996; 96US-00739167.
XX
PR 15-SEP-1993; 93US-00122791.
PR 18-FEB-1994; 94US-00198450.

PR 30-NOV-1994; 94US-00348472.
PR 20-JAN-1995; 95US-00376184.
PR 15-MAR-1995; 95US-00404796.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chang SMW, Jolly DJ, Dubensky TW, Belli BA, Ibanez CE, Driver DA;
PI Polo JM;
XX
DR WPI; 1999-044581/04.
XX
PT Alphavirus vectors constructs containing a 5' promoter of viral cDNA by
PT in vitro transcription - used in gene therapy.
XX
PS Example 7; Col 103; 140pp; English.
XX
CC PCR primers AAV70745-46 are used to amplify part of the hepatitis delta
CC virus (HDV) genome. The amplified product is used in the production of
CC the alphavirus vector constructs of the invention. These constructs
CC comprise a promoter 5' of viral cDNA which initiates the synthesis of RNA
CC from the viral cDNA by in vitro transcription, followed by a 5' sequence
CC which initiates transcription of alphavirus RNA, followed by a nucleotide
CC sequence encoding alphavirus nonstructural proteins, a viral junction
CC region which has been inactivated such that viral transcription of a
CC subgenomic fragment is prevented, an internal ribosome entry site or a
CC sequence which promotes ribosome read through between adjacent reading
CC frames, and an alphavirus RNA polymerase recognition sequence. The
CC recombinant alphavirus vectors can be used for gene therapy. (Updated on
CC 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 51.9%; Score 29.6; DB 2; Length 52;
Best Local Similarity 59.4%; Pred. No. 43;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32
```

RESULT 15
AAV70714
ID AAV70714 standard; DNA; 52 BP.
XX
AC AAV70714;
XX
DT 20-MAR-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Nested primer HDV17-68 for hepatitis delta virus ribozyme sequence.
XX
KW Alphavirus vector construct; gene therapy; PCR primer; ss.
XX
OS Synthetic.
OS Hepatitis D virus.
XX
PN US5843723-A.
XX
PD 01-DEC-1998.
XX
PF 30-OCT-1996; 96US-00739167.
XX
PR 15-SEP-1993; 93US-00122791.
PR 18-FEB-1994; 94US-00198450.
PR 30-NOV-1994; 94US-00348472.
PR 20-JAN-1995; 95US-00376184.
PR 15-MAR-1995; 95US-00404796.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chang SMW, Jolly DJ, Dubensky TW, Belli BA, Ibanez CE, Driver DA;
PI Polo JM;
XX
DR WPI; 1999-044581/04.
XX
PT Alphavirus vectors constructs containing a 5' promoter of viral cDNA by
PT in vitro transcription - used in gene therapy.
XX
PS Example 3; Col 69; 140pp; English.
XX
CC PCR primers AAV70713-15 are overlapping primers used to amplify hepatitis
CC delta virus (HDV) ribozyme sequences. The amplified product is used in
CC the production of the alphavirus vector constructs of the invention.
CC These constructs comprise a promoter 5' of viral cDNA which initiates the
CC synthesis of RNA from the viral cDNA by in vitro transcription, followed
CC by a 5' sequence which initiates transcription of alphavirus RNA,
CC followed by a nucleotide sequence encoding alphavirus nonstructural
CC proteins, a viral junction region which has been inactivated such that
CC viral transcription of a subgenomic fragment is prevented, an internal
CC ribosome entry site or a sequence which promotes ribosome read through
CC between adjacent reading frames, and an alphavirus RNA polymerase
CC recognition sequence. The recombinant alphavirus vectors can be used for
CC gene therapy. (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 52 BP; 9 A; 21 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 51.9%; Score 29.6; DB 2; Length 52;
Best Local Similarity 59.4%; Pred. No. 43;
Matches 19; Conservative 12; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32
```

Search completed: May 19, 2004, 03:06:11
Job time : 407 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 19, 2004, 01:36:25 ; Search time 3047 Seconds
(without alignments)
810.815 Million cell updates/sec

Title: RIBOZYME1A
Perfect score: 57
Sequence: 1 ggguccaccuccucgcggun.....uucgcauggcuaagggaccc 57

Scoring table: IDENTITY_NUCDX
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		56.2	98.6	57	6	AX012282	AX012282 Sequence
2		56.2	98.6	57	6	AX012283	AX012283 Sequence
3		56.2	98.6	57	6	AX012286	AX012286 Sequence
4		56.2	98.6	57	6	AX012287	AX012287 Sequence
5		56.2	98.6	57	6	AX012288	AX012288 Sequence
6		56.2	98.6	57	6	AX012289	AX012289 Sequence
7		56.2	98.6	57	6	AX012290	AX012290 Sequence
8		55.4	97.2	57	6	AX012285	AX012285 Sequence
9		46.6	81.8	57	6	AX012284	AX012284 Sequence
10		36.2	63.5	37	6	AX012291	AX012291 Sequence
11	c	31.4	55.1	168545	9	HSA287B20	AL121896 Human DNA
12		31	54.4	61	6	I41390	I41390 Sequence 18
13	c	29.8	52.3	159423	9	AC008429	AC008429 Homo sapi
14		29.8	52.3	171914	2	AC119266	AC119266 Mus muscu
15		29.8	52.3	184032	2	AC117632	AC117632 Mus muscu
16	c	29.8	52.3	213025	2	AC021159	AC021159 Homo sapi
17		29.6	51.9	52	6	AR021007	AR021007 Sequence
18		29.6	51.9	52	6	AR021047	AR021047 Sequence
19		29.6	51.9	52	6	AR043422	AR043422 Sequence
20		29.6	51.9	52	6	AR043462	AR043462 Sequence
21		29.6	51.9	52	6	AR062337	AR062337 Sequence
22		29.6	51.9	52	6	AR062377	AR062377 Sequence
23		29.6	51.9	52	6	AR183796	AR183796 Sequence
24		29.6	51.9	52	6	AR183836	AR183836 Sequence
25		29.6	51.9	52	6	AR368179	AR368179 Sequence
26		29.6	51.9	52	6	AR368219	AR368219 Sequence
27		29.6	51.9	80	6	I41388	I41388 Sequence 16
28		29.6	51.9	83	6	AR123021	AR123021 Sequence
29		29.6	51.9	84	6	I41385	I41385 Sequence 13
30		29.6	51.9	84	6	I41386	I41386 Sequence 14
31		29.6	51.9	86	6	I41378	I41378 Sequence 6
32		29.6	51.9	90	6	AR243300	AR243300 Sequence
33		29.6	51.9	90	6	AR342596	AR342596 Sequence
34	c	29.6	51.9	90	6	AR342597	AR342597 Sequence
35		29.6	51.9	90	6	AX174838	AX174838 Sequence
36	c	29.6	51.9	90	6	AX174839	AX174839 Sequence
37		29.6	51.9	90	6	BD226817	BD226817 Alphaviru
38		29.6	51.9	101	6	AR363756	AR363756 Sequence
39	c	29.6	51.9	105	6	A41844	A41844 Sequence 25
40	c	29.6	51.9	105	6	A41849	A41849 Sequence 30
41		29.6	51.9	567	6	I05156	I05156 Sequence 11
42	c	29.6	51.9	866	14	HPDJS1	D90192 Hepatitis D
43	c	29.6	51.9	866	14	HPDJS2	D90193 Hepatitis D
44	c	29.6	51.9	868	14	HPDJM1	D90190 Hepatitis D
45	c	29.6	51.9	868	14	HPDJM2	D90191 Hepatitis D

ALIGNMENTS

RESULT 1
AX012282
LOCUS AX012282 57 bp RNA linear PAT 06-SEP-2000
DEFINITION Sequence 44 from Patent WO9955856.
ACCESSION AX012282
VERSION AX012282.1 GI:9998343
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
TITLE Nucleic acid enzyme for rna cleavage
JOURNAL Patent: WO 9955856-A 44 04-NOV-1999;
ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN

```
           PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGACCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 2
AX012283
LOCUS       AX012283                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 45 from Patent WO9955856.
ACCESSION   AX012283
VERSION     AX012283.1  GI:9998344
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 45 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCCAGCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 3
AX012286
LOCUS       AX012286                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 48 from Patent WO9955856.
ACCESSION   AX012286
VERSION     AX012286.1  GI:9998347
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 48 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGTCCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 4
AX012287
LOCUS       AX012287                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 49 from Patent WO9955856.
ACCESSION   AX012287
VERSION     AX012287.1  GI:9998348
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 49 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGGCCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 5
AX012288
LOCUS       AX012288                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 50 from Patent WO9955856.
ACCESSION   AX012288
VERSION     AX012288.1  GI:9998349
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 50 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
```

```
                ||||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGAACTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 6
AX012289
LOCUS       AX012289                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 51 from Patent WO9955856.
ACCESSION   AX012289
VERSION     AX012289.1  GI:9998350
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 51 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGAGCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 7
AX012290
LOCUS       AX012290                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 52 from Patent WO9955856.
ACCESSION   AX012290
VERSION     AX012290.1  GI:9998351
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 52 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             98.6%;  Score 56.2;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 1.3e-05;
  Matches   41;  Conservative   16;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGATCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 8
AX012285
LOCUS       AX012285                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 47 from Patent WO9955856.
ACCESSION   AX012285
VERSION     AX012285.1  GI:9998346
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 47 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             97.2%;  Score 55.4;  DB 6;  Length 57;
  Best Local Similarity   71.9%;  Pred. No. 2.3e-05;
  Matches   41;  Conservative   15;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||:::: :::|||||:|||||::||||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGCCCTGGGCATGCGGCTTCGCATGGCTAAGGGACCC 57


RESULT 9
AX012284
LOCUS       AX012284                 57 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 46 from Patent WO9955856.
ACCESSION   AX012284
VERSION     AX012284.1  GI:9998345
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 46 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             81.8%;  Score 46.6;  DB 6;  Length 57;
  Best Local Similarity   64.9%;  Pred. No. 0.012;
  Matches   37;  Conservative   14;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||:|||||:||:|||||::::::::|||||:||       |||:|||:|||||||||
Db          1 GGGTCCACCTCCTCGCGGTCCGACCTGGGCATGCCTTCGGGCATGGCTAAGGGACCC 57


RESULT 10
AX012291
LOCUS       AX012291                 37 bp    RNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 53 from Patent WO9955856.
ACCESSION   AX012291
VERSION     AX012291.1  GI:9998352
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
```

```
REFERENCE   1
  AUTHORS   Ananvoranich,S., Lafontaine,D. and Perreault,J.P.
  TITLE     Nucleic acid enzyme for rna cleavage
  JOURNAL   Patent: WO 9955856-A 53 04-NOV-1999;
            ANANVORANICH SIRINART (CA); LAFONTAINE DANIEL (CA); PERREAULT JEAN
            PIERRE (CA); UNIV SHERBROOKE (CA)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid"
ORIGIN

  Query Match             63.5%;  Score 36.2;  DB 6;  Length 37;
  Best Local Similarity   67.6%;  Pred. No. 21;
  Matches   25;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGC 37
              |||:|||||:||:|||||::::::::|||||:|||||
Db          1 GGGTCCACCTCCTCGCGGTCCGACCTGGGCATGCGGC 37


RESULT 11
HSA287B20/c
LOCUS       HSA287B20             168545 bp    DNA     linear   PRI 11-MAR-2001
DEFINITION  Human DNA sequence from clone RP11-287B20 on chromosome 20 Contains
            part of gene KIAA1272 for a protein similar to rat Tulip 2, ESTs,
            STSs and GSSs, complete sequence.
ACCESSION   AL121896
VERSION     AL121896.11  GI:12583641
KEYWORDS    HTG; KIAA1272; Tulip.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 168545)
  AUTHORS   Griffiths,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 27, 2001 this sequence version replaced gi:8218066.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            RP11-287B20 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-287B20 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-287B20 is at 1 in this sequence.
            The true left end of clone RP5-1049G11 is at 168446 in this
            sequence. The true right end of clone RP11-470C13 is at 111612 in
            this sequence. This sequence was finished as follows unless
            otherwise noted: all regions were either double-stranded or
            sequenced with an alternate chemistry or covered by high quality
            data (i.e., phred quality >= 30); an attempt was made to resolve
            all sequencing problems, such as compressions and repeats; all
            regions were covered by at least one plasmid subclone or more than
            one M13 subclone; and the assembly was confirmed by restriction
            digest.
FEATURES             Location/Qualifiers
     source          1. .168545
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP11-287B20"
                     /clone_lib="RPCI-11.1"
     misc_feature    23. .649
                     /note="match: GSS: Em:AQ542006"
     repeat_region   30. .496
                     /note="L1MB8 repeat: matches 5689. .6167 of consensus"
     misc_feature    complement(358. .678)
                     /note="match: STS: Em:G25753 Em:G27399"
     repeat_region   2049. .2584
                     /note="L2 repeat: matches 2101. .2670 of consensus"
     repeat_region   5422. .5607
                     /note="MER5B repeat: matches 1. .178 of consensus"
     repeat_region   5931. .6206
                     /note="AluJb repeat: matches 1. .275 of consensus"
     misc_feature    complement(7762. .8243)
                     /note="match: GSS: Em:AZ260027"
     misc_feature    9279. .9897
                     /note="match: STS: Em:G57973
                     match: GSS: Em:AQ345124"
     misc_feature    9283. .9746
                     /note="match: GSS: Em:AQ525450"
     misc_feature    complement(9548. .10038)
                     /note="match: GSS: Em:AQ772919"
     misc_feature    10005. .10404
                     /note="match: GSS: Em:AQ409384"
     gene            complement(11328. .146133)
                     /gene="bA287B20.1"
     CDS             complement(join(<11328. .11449,72130. .72257,94418. .94576,
                     102652. .102807,104712. .104835,110555. .110626,
                     111814. .112585,120218. .120401,123707. .123844,
                     124131. .124199,125553. .125663,130988. .131096,
                     135046. .135104,135957. .136085,146005. .>146133))
                     /gene="bA287B20.1"
                     /note="continued from dJ1049G11 (AL078634)
                     match: cDNAs: Em:AK002211 Em:AB032400 Em:AF041106
                     Em:AF041107 Em:AL050050 Em:AB033098
                     match: ESTs: Em:T90823 Em:BE467241 Em:BE349987 Em:BE219600
                     Em:R05337 Em:T85728 Em:AW753998 Em:AW604552 Em:AW658646
                     Em:R05338
                     match: proteins: Tr:O55008 Tr:Q9Y408"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="bA287B20.1.1 (KIAA1272 similar to rat tulip
                     proteins 1 and 2, isoform 1)"
                     /protein_id="CAC27349.1"
                     /db_xref="GI:12583643"
                     /db_xref="SPTREMBL:Q9BQT6"
                     /translation="DILNTIIRHCPPRFFSLGFPGFSMLVGDFITAAARVLSTDILTA
                     PRSEAVTVLGSLVCFPNTYQEIPLLQSVPEVNEAITGTEDVKHYLINILLKNATEEPN
                     EYARCIAVCSLGVWICEELAQCTSHPQVKEAINVIGVTLKFPNKIVAQVACDVLQLLV
                     SYWEKLQMFETSLPRKMAEILVATVAFLLPSAEYSSVETDKKFIVSLLLCLLDWCMAL
                     PVSVLLHPVSTAVLEEQHSARAPLLDYIYRVLHCCVCGSSTYTQQSHYILTLADLSST
                     DYDPFLPLANVKSSEPVQYHSSAELGNLLTVEEEKKRRSLELIPLTARMVMAHLVNHL
                     GHYPLSGGPAILHSLVSENHDNAHVEGSELSFEVFRSPNLQLFVFNDSTLISYLQTPT
                     EGPVGGSPVGSLSDVRVIVRDISGKYSWDGKVLYGPLEGCLAPNGRNPSFLISSWHRD
                     TFGPQKDSSQVEEGDDVLDKLLENIGHTSPECLLPSQLNLNEPSLTPCGMNYDQEKEI
                     IEVILRQNAQEDEYIQSHNFDSAMKVTSQGQPSPVEPRGPFYFCRLLLDDLGMNSWDR
                     RKNFHLLKKNSKLLRELKNLDSRQCRETHKIAVFYIAEGQEDKCSILSNERGSQAYED
                     FVAGLGWEVDLSTHCGFMGGLQRNGSTGQTAPYYATSTVEVIFHVSTRMPSDSDDSLT
                     KKLRHLGNDEVHIVWSEHSRDYRRGIIPTAFGDVSIIIYPMKNHMFFIAITKKPEVPF
                     FGPLFDGAIVSGKLLPSLVCATCINASRAVKCLIPLYQSFYEERALYLEAIIQNHREV
                     MTFEDFAAQVFSPSPSYSLSGT"
     misc_feature    11547. .12050
                     /note="match: STS: Em:G55717
```

```
                match: GSS: Em:AQ312069"
repeat_region   12766. .12980
                /note="MER3 repeat: matches 4. .209 of consensus"
repeat_region   14926. .15041
                /note="FLAM_A repeat: matches 20. .135 of consensus"
repeat_region   15330. .15369
                /note="MIR repeat: matches 106. .148 of consensus"
repeat_region   15335. .15474
                /note="MIR repeat: matches 70. .213 of consensus"
repeat_region   15634. .15733
                /note="MIR repeat: matches 81. .175 of consensus"
misc_feature    16054. .16481
                /note="match: GSS: Em:B47002"
repeat_region   16546. .16696
                /note="MLT1C repeat: matches 299. .466 of consensus"
repeat_region   16740. .16951
                /note="MLT1C repeat: matches 17. .237 of consensus"
repeat_region   17228. .17518
                /note="L1M4 repeat: matches 5031. .5329 of consensus"
repeat_region   17900. .18040
                /note="L1MD2 repeat: matches 5788. .5932 of consensus"
repeat_region   18042. .18340
                /note="AluJo repeat: matches 1. .298 of consensus"
repeat_region   19137. .19367
                /note="L2 repeat: matches 2515. .2748 of consensus"
repeat_region   20304. .20689
                /note="L1MB7 repeat: matches 5188. .5577 of consensus"
repeat_region   20690. .21000
                /note="AluSp repeat: matches 1. .312 of consensus"
repeat_region   21001. .21605
                /note="L1MB7 repeat: matches 5577. .6169 of consensus"
repeat_region   21618. .21881
                /note="AluJo repeat: matches 1. .292 of consensus"
repeat_region   22065. .22161
                /note="L2 repeat: matches 2577. .2716 of consensus"
repeat_region   23992. .24102
                /note="L1PA4 repeat: matches 6032. .6142 of consensus"
repeat_region   26203. .26306
                /note="MIR repeat: matches 102. .208 of consensus"
repeat_region   28013. .28104
                /note="23 copies 4 mer ccac 80% conserved"
repeat_region   28439. .28570
                /note="MIR repeat: matches 1. .136 of consensus"
repeat_region   29293. .29412
                /note="MIR repeat: matches 23. .147 of consensus"
repeat_region   30933. .31124
                /note="L1MA7 repeat: matches 6075. .6273 of consensus"
misc_feature    32304. .32730
                /note="match: GSS: Em:AQ176627"
repeat_region   32443. .32571
                /note="MIR repeat: matches 100. .237 of consensus"
repeat_region   33955. .34242
                /note="AluJb repeat: matches 1. .288 of consensus"
repeat_region   36573. .37174
                /note="L2 repeat: matches 1921. .2661 of consensus"
repeat_region   37353. .37774
                /note="Charlie4a repeat: matches 30. .501 of consensus"
repeat_region   38519. .38813
                /note="AluJb repeat: matches 1. .295 of consensus"
repeat_region   41378. .41532
                /note="MIR repeat: matches 1. .174 of consensus"
misc_feature    41535. .41666
                /note="match: GSS: Em:AQ626507"
repeat_region   43258. .43426
                /note="MER5B repeat: matches 1. .175 of consensus"
repeat_region   43745. .43776
                /note="8 copies 4 mer acac 87% conserved"
repeat_region   43751. .43776
                /note="13 copies 2 mer ac 92% conserved"
repeat_region   43783. .43862
                /note="L1ME3A repeat: matches 6030. .6117 of consensus"
repeat_region   43944. .44059
                /note="L1ME repeat: matches 5469. .5587 of consensus"
repeat_region   44070. .44541
                /note="MER44B repeat: matches 1. .533 of consensus"
repeat_region   45723. .45772
                /note="MADE1 repeat: matches 31. .80 of consensus"
repeat_region   45848. .45909
                /note="MIR repeat: matches 148. .207 of consensus"

  Query Match             55.1%;  Score 31.4;  DB 9;  Length 168545;
  Best Local Similarity   50.9%;  Pred. No. 2e+02;
  Matches   27;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 CCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              ||||  ||:|  ||::::::::|||||:|| ||::||   |||:  | |  ||
Db       2962 CCACAGCCTCCAGGTAACATCTGGGCATGCAGCTTCGTGAGGCTGGGCGCTCC 2910


RESULT 12
I41390
LOCUS       I41390                    61 bp    DNA     linear   PAT 13-MAY-1997
DEFINITION  Sequence 18 from patent US 5625047.
ACCESSION   I41390
VERSION     I41390.1  GI:2081980
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Been,M.D., Rosenstein,S.P. and Perrotta,A.T.
  TITLE     Enzymatic RNA molecules
  JOURNAL   Patent: US 5625047-A 18 29-APR-1997;
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             54.4%;  Score 31;  DB 6;  Length 61;
  Best Local Similarity   58.5%;  Pred. No. 8e+02;
  Matches   31;  Conservative   16;  Mismatches    2;  Indels    4;  Gaps    1;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACC 56
              :|||||:||:|||||::::::::|||||:    |::|| |:|||:|||||| |
Db          8 TCCACCTCCTCGCGGTCCGACCTGGGCAT----CTTCGGATGGCTAAGGGAGC 56


RESULT 13
AC008429/c
LOCUS       AC008429               159423 bp    DNA     linear   PRI 05-NOV-2002
DEFINITION  Homo sapiens chromosome 5 clone CTC-308K20, complete sequence.
ACCESSION   AC008429
VERSION     AC008429.6  GI:24580340
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 159423)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 159423)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 159423)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Nov 5, 2002 this sequence version replaced gi:15290285.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
```

```
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
            NOTE: Shatter libraries failed to resolve dinucleotide repeat
            region from 21938 to 21990. Forced join at 21980.
FEATURES              Location/Qualifiers
     source           1. .159423
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="5"
                      /clone="CTC-308K20"
     misc_feature     21938. .21990
                      /note="Shatter libraries failed to resolve dinucleotide
                      repeat region from 21938 to 21990. Forced join at 21980."
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 9;  Length 159423;
  Best Local Similarity   49.1%;  Pred. No. 6.4e+02;
  Matches   26;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          5 CCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGACCC 57
              |||||:| : || | :::::::|||||  | |||::| ||:   : ||   |||
Db      17452 CCACCTCTTGGCTGATTGACTTGGGCAAGGGGCTTCCCATCCATCAGTTTCCC 17400


RESULT 14
AC119266
LOCUS       AC119266              171914 bp    DNA     linear   HTG 17-MAR-2003
DEFINITION  Mus musculus clone RP24-321G8, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION   AC119266
VERSION     AC119266.3  GI:28975955
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 171914)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-321G8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171914)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 171914)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 17, 2003 this sequence version replaced gi:25956367.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L25499
                Center clone name: 321_G_8
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 169262 bases at least Q40
                Consensus quality: 170279 bases at least Q30
                Consensus quality: 170654 bases at least Q20
                Insert size: 176000; agarose-fp
                Insert size: 170914; sum-of-contigs
                Quality coverage: 10.1 in Q20 bases; agarose-fp
                Quality coverage: 10.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    13872: contig of 13872 bp in length
            *    13873    13972: gap of 100 bp
            *    13973    14646: contig of 674 bp in length
            *    14647    14746: gap of 100 bp
            *    14747    16119: contig of 1373 bp in length
            *    16120    16219: gap of 100 bp
            *    16220    17367: contig of 1148 bp in length
            *    17368    17467: gap of 100 bp
            *    17468    19014: contig of 1547 bp in length
            *    19015    19114: gap of 100 bp
            *    19115    31449: contig of 12335 bp in length
            *    31450    31549: gap of 100 bp
            *    31550    52958: contig of 21409 bp in length
            *    52959    53058: gap of 100 bp
            *    53059    70073: contig of 17015 bp in length
            *    70074    70173: gap of 100 bp
            *    70174    97112: contig of 26939 bp in length
            *    97113    97212: gap of 100 bp
            *    97213   132535: contig of 35323 bp in length
            *   132536   132635: gap of 100 bp
            *   132636   171914: contig of 39279 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1. .171914
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP24-321G8"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature    1. .13872
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    13973. .14646
                     /note="assembly_fragment"
     misc_feature    14747. .16119
                     /note="assembly_fragment"
     misc_feature    16220. .17367
                     /note="assembly_fragment"
     misc_feature    17468. .19014
                     /note="assembly_fragment"
     misc_feature    19115. .31449
                     /note="assembly_fragment"
     misc_feature    31550. .52958
                     /note="assembly_fragment"
     misc_feature    53059. .70073
                     /note="assembly_fragment"
     misc_feature    70174. .97112
                     /note="assembly_fragment"
     misc_feature    97213. .132535
                     /note="assembly_fragment"
     misc_feature    132636. .171914
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 2;  Length 171914;
  Best Local Similarity   45.3%;  Pred. No. 6.3e+02;
  Matches   24;  Conservative   15;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          2 GGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGA 54
              ||: || |:||: | ||::::::::|||   | || ::| | :|| :| || |
Db     106817 GGTGCAGCTCCTTGAGGTGGGGGCTGGGGGCGGGGTTTCTCCTGGGTAGGGTA 106869


RESULT 15
AC117632
LOCUS       AC117632               184032 bp    DNA     linear   HTG 08-FEB-2003
DEFINITION  Mus musculus clone RP23-187D8, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION   AC117632
VERSION     AC117632.3  GI:28273439
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 184032)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-187D8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184032)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 184032)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 8, 2003 this sequence version replaced gi:24962351.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L23511
                Center clone name: 187_D_8
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 181445 bases at least Q40
                Consensus quality: 182384 bases at least Q30
                Consensus quality: 182696 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 183032; sum-of-contigs
                Quality coverage: 10.8 in Q20 bases; agarose-fp
                Quality coverage: 10.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    24767: contig of 24767 bp in length
            *    24768    24867: gap of 100 bp
            *    24868    26041: contig of 1174 bp in length
            *    26042    26141: gap of 100 bp
```

```
            *    26142    29590: contig of 3449 bp in length
            *    29591    29690: gap of 100 bp
            *    29691    31311: contig of 1621 bp in length
            *    31312    31411: gap of 100 bp
            *    31412    41665: contig of 10254 bp in length
            *    41666    41765: gap of 100 bp
            *    41766    54699: contig of 12934 bp in length
            *    54700    54799: gap of 100 bp
            *    54800    84383: contig of 29584 bp in length
            *    84384    84483: gap of 100 bp
            *    84484   112216: contig of 27733 bp in length
            *   112217   112316: gap of 100 bp
            *   112317   140050: contig of 27734 bp in length
            *   140051   140150: gap of 100 bp
            *   140151   175535: contig of 35385 bp in length
            *   175536   175635: gap of 100 bp
            *   175636   184032: contig of 8397 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184032
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-187D8"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .24767
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    24868. .26041
                     /note="assembly_fragment"
     misc_feature    26142. .29590
                     /note="assembly_fragment"
     misc_feature    29691. .31311
                     /note="assembly_fragment"
     misc_feature    31412. .41665
                     /note="assembly_fragment"
     misc_feature    41766. .54699
                     /note="assembly_fragment"
     misc_feature    54800. .84383
                     /note="assembly_fragment"
     misc_feature    84484. .112216
                     /note="assembly_fragment"
     misc_feature    112317. .140050
                     /note="assembly_fragment"
     misc_feature    140151. .175535
                     /note="assembly_fragment"
     misc_feature    175636. .184032
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             52.3%;  Score 29.8;  DB 2;  Length 184032;
  Best Local Similarity   45.3%;  Pred. No. 6.3e+02;
  Matches   24;  Conservative   15;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          2 GGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGGCUAAGGGA 54
              ||: || |:||: | ||::::::::|||   | || ::| | :|| :| || |
Db     149817 GGTGCAGCTCCTTGAGGTGGGGGCTGGGGGCGGGGTTTCTCCTGGGTAGGGTA 149869


Search completed: May 19, 2004, 03:57:14
Job time : 3052 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 19, 2004, 02:59:41 ; Search time 390 Seconds
(without alignments)
663.253 Million cell updates/sec

Title: RIBOZYME1A
Perfect score: 57
Sequence: 1 ggguccaccuccucgcggun.....uucgcauggcuaagggaccc 57

Scoring table: IDENTITY_NUCDX
Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	29.6	51.9	52	13	US-10-150-407-40	Sequence 40, Appl
	2	29.6	51.9	52	13	US-10-150-407-80	Sequence 80, Appl
	3	29.6	51.9	52	16	US-10-346-880-40	Sequence 40, Appl
	4	29.6	51.9	52	16	US-10-346-880-80	Sequence 80, Appl
	5	29.6	51.9	90	9	US-09-733-042-12	Sequence 12, Appl
c	6	29.6	51.9	90	9	US-09-733-042-13	Sequence 13, Appl
	7	29.6	51.9	90	10	US-09-190-246-2	Sequence 2, Appli
	8	29.6	51.9	91	15	US-10-237-302-7	Sequence 7, Appli
c	9	29.6	51.9	99	15	US-10-237-302-8	Sequence 8, Appli
	10	29.6	51.9	8100	10	US-09-190-246-4	Sequence 4, Appli
	11	29.6	51.9	15538	10	US-09-190-246-1	Sequence 1, Appli
c	12	29.2	51.2	466	10	US-09-918-995-2716	Sequence 2716, Ap
c	13	29.2	51.2	487	10	US-09-918-995-10300	Sequence 10300, A
c	14	29.2	51.2	601	14	US-10-109-860-6	Sequence 6, Appli
c	15	29.2	51.2	1415	13	US-10-342-887-515	Sequence 515, App
c	16	29.2	51.2	1415	13	US-10-172-118-515	Sequence 515, App
c	17	29.2	51.2	1445	14	US-10-109-860-1	Sequence 1, Appli
c	18	29.2	51.2	9704	14	US-10-109-860-3	Sequence 3, Appli
c	19	29	50.9	564	15	US-10-029-386-4237	Sequence 4237, Ap
	20	28.8	50.5	2730	15	US-10-156-761-178	Sequence 178, App
c	21	28.8	50.5	3300	16	US-10-260-238-1175	Sequence 1175, Ap
	22	28.8	50.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
c	23	28.4	49.8	3386	15	US-10-101-510-156	Sequence 156, App
c	24	28.4	49.8	3386	16	US-10-116-275-209	Sequence 209, App
c	25	28.4	49.8	3386	17	US-10-651-836-7	Sequence 7, Appli
	26	28.2	49.5	56	10	US-09-507-362-39	Sequence 39, Appl
	27	28.2	49.5	56	16	US-10-391-441-39	Sequence 39, Appl
	28	28.2	49.5	1170	13	US-10-047-825-32	Sequence 32, Appl
c	29	28	49.1	372	13	US-10-085-783A-55374	Sequence 55374, A
c	30	28	49.1	372	16	US-10-242-535A-55374	Sequence 55374, A
c	31	28	49.1	1691139	15	US-10-067-514-1	Sequence 1, Appli
c	32	28	49.1	1691139	16	US-10-419-723-1	Sequence 1, Appli
c	33	27.8	48.8	842	15	US-10-252-157-493	Sequence 493, App
c	34	27.8	48.8	4711	15	US-10-084-817-183	Sequence 183, App
c	35	27.8	48.8	4732	15	US-10-133-013-17	Sequence 17, Appl
	36	27.8	48.8	11009	9	US-09-845-583-1	Sequence 1, Appli
	37	27.8	48.8	11009	15	US-10-037-182-3	Sequence 3, Appli
	38	27.4	48.1	1519	13	US-10-047-825-7	Sequence 7, Appli
c	39	27.4	48.1	24081	15	US-10-132-134-13	Sequence 13, Appl
c	40	27.4	48.1	52101	15	US-10-132-134-1	Sequence 1, Appli
c	41	27.2	47.7	897	9	US-09-771-161A-26	Sequence 26, Appl
c	42	27.2	47.7	1200	16	US-10-369-493-24358	Sequence 24358, A
	43	27.2	47.7	1218	13	US-10-282-122A-30258	Sequence 30258, A
c	44	27.2	47.7	1679	9	US-09-771-161A-25	Sequence 25, Appl
	45	27.2	47.7	1956	17	US-10-416-314-101	Sequence 101, App

ALIGNMENTS

RESULT 1
US-10-150-407-40
; Sequence 40, Application US/10150407
; Publication No. US20040029278A1
; GENERAL INFORMATION:
; APPLICANT: Dubensky Jr, Thomas W
; Polo, John M.
; Jolly, Douglas J.
; Driver, David A.
; TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/150,407
; FILING DATE: 17-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/350,522
; FILING DATE: 08-Jul-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

```
;   INFORMATION FOR SEQ ID NO: 40:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-150-407-40

  Query Match             51.9%;  Score 29.6;  DB 13;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 21;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 2
US-10-150-407-80
; Sequence 80, Application US/10150407
; Publication No. US20040029278A1
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr, Thomas W
;                   Polo, John M.
;                   Jolly, Douglas J.
;                   Driver, David A.
;        TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;        NUMBER OF SEQUENCES: 128
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: US
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/150,407
;             FILING DATE: 17-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,522
;             FILING DATE: 08-Jul-1999
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 80:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-10-150-407-80

  Query Match             51.9%;  Score 29.6;  DB 13;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 21;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 3
US-10-346-880-40
; Sequence 40, Application US/10346880
; Publication No. US20030232035A1
; GENERAL INFORMATION:
;  APPLICANT: Dubensky Jr., Thomas W
;  APPLICANT:  Polo, John M
;  APPLICANT:  Ibanez, Carlos E
;  APPLICANT:  Chang, Stephen M.W.
;  APPLICANT:  Jolly, Douglas J
;  APPLICANT:  Driver, David A
;  APPLICANT:  Belli, Barbara A
;  TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;  FILE REFERENCE: PP01146.213 (CHIR-1146/11US)
;  CURRENT APPLICATION NUMBER: US/10/346,880
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 09/503,138
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: 09/191,747
;  PRIOR FILING DATE: 1998-11-12
;  PRIOR APPLICATION NUMBER: 08/739,199
;  PRIOR FILING DATE: 1996-10-30
;  PRIOR APPLICATION NUMBER: 08/404,796
;  PRIOR FILING DATE: 1995-03-15
;  PRIOR APPLICATION NUMBER: 08/376,184
;  PRIOR FILING DATE: 1995-01-18
;  PRIOR APPLICATION NUMBER: 08/348,472
;  PRIOR FILING DATE: 1994-11-30
;  PRIOR APPLICATION NUMBER: 08/198,450
;  PRIOR FILING DATE: 1994-02-18
;  PRIOR APPLICATION NUMBER: 08/122,791
;  PRIOR FILING DATE: 1993-09-15
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 40
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Artificial sequence: nested primer HDV17-68
US-10-346-880-40

  Query Match             51.9%;  Score 29.6;  DB 16;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 21;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 4
US-10-346-880-80
; Sequence 80, Application US/10346880
; Publication No. US20030232035A1
; GENERAL INFORMATION:
;  APPLICANT: Dubensky Jr., Thomas W
;  APPLICANT:  Polo, John M
;  APPLICANT:  Ibanez, Carlos E
;  APPLICANT:  Chang, Stephen M.W.
;  APPLICANT:  Jolly, Douglas J
;  APPLICANT:  Driver, David A
;  APPLICANT:  Belli, Barbara A
;  TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;  FILE REFERENCE: PP01146.213 (CHIR-1146/11US)
;  CURRENT APPLICATION NUMBER: US/10/346,880
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 09/503,138
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: 09/191,747
;  PRIOR FILING DATE: 1998-11-12
;  PRIOR APPLICATION NUMBER: 08/739,199
;  PRIOR FILING DATE: 1996-10-30
```

```
;  PRIOR APPLICATION NUMBER: 08/404,796
;  PRIOR FILING DATE: 1995-03-15
;  PRIOR APPLICATION NUMBER: 08/376,184
;  PRIOR FILING DATE: 1995-01-18
;  PRIOR APPLICATION NUMBER: 08/348,472
;  PRIOR FILING DATE: 1994-11-30
;  PRIOR APPLICATION NUMBER: 08/198,450
;  PRIOR FILING DATE: 1994-02-18
;  PRIOR APPLICATION NUMBER: 08/122,791
;  PRIOR FILING DATE: 1993-09-15
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 80
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Artificial sequence: reverse primer HKV17-68
US-10-346-880-80

  Query Match             51.9%;  Score 29.6;  DB 16;  Length 52;
  Best Local Similarity   59.4%;  Pred. No. 21;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db          1 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 32


RESULT 5
US-09-733-042-12
; Sequence 12, Application US/09733042
; Patent No. US20020168709A1
; GENERAL INFORMATION:
;  APPLICANT:   Hennecke, Frank
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION:   Replicon Based Activation of Endogenous Genes
;  FILE REFERENCE:   1700.0100001
;  CURRENT APPLICATION NUMBER: US/09/733,042
;  CURRENT FILING DATE:  2000-12-11
;  PRIOR APPLICATION NUMBER:   US 60/169,988
;  PRIOR FILING DATE:   1999-12-10
;  NUMBER OF SEQ ID NOS:   49
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: HDV-FOR
US-09-733-042-12

  Query Match             51.9%;  Score 29.6;  DB 9;  Length 90;
  Best Local Similarity   59.4%;  Pred. No. 23;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db         22 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 53


RESULT 6
US-09-733-042-13/c
; Sequence 13, Application US/09733042
; Patent No. US20020168709A1
; GENERAL INFORMATION:
;  APPLICANT:   Hennecke, Frank
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION:   Replicon Based Activation of Endogenous Genes
;  FILE REFERENCE:   1700.0100001
;  CURRENT APPLICATION NUMBER: US/09/733,042
;  CURRENT FILING DATE:  2000-12-11
;  PRIOR APPLICATION NUMBER:   US 60/169,988
;  PRIOR FILING DATE:   1999-12-10
;  NUMBER OF SEQ ID NOS:   49
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: HDV-REV
US-09-733-042-13

  Query Match             51.9%;  Score 29.6;  DB 9;  Length 90;
  Best Local Similarity   59.4%;  Pred. No. 23;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db         73 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 42


RESULT 7
US-09-190-246-2
; Sequence 2, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT: Li, Xiaomao
;  APPLICANT: Klein, Michel
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: Parrington et al.
;  CURRENT APPLICATION NUMBER: US/09/190,246
;  CURRENT FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Semliki Forest virus
US-09-190-246-2

  Query Match             51.9%;  Score 29.6;  DB 10;  Length 90;
  Best Local Similarity   59.4%;  Pred. No. 23;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db         22 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 53


RESULT 8
US-10-237-302-7
; Sequence 7, Application US/10237302
; Publication No. US20030119182A1
; GENERAL INFORMATION:
;  APPLICANT: AlphaVax, Inc.
;  APPLICANT:  Jonathan F. Smith
;  APPLICANT:  Kurt I. Kamrud
;  APPLICANT:  Jonathan O. Rayner
;  APPLICANT:  Sergey A. Dryga
;  APPLICANT:  Ian J. Caley
;  TITLE OF INVENTION: ALPHAVIRUS REPLICON VECTOR SYSTEMS
;  FILE REFERENCE: 01113.0002U2
;  CURRENT APPLICATION NUMBER: US/10/237,302
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 60/317,722
;  PRIOR FILING DATE: 2001-09-06
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence; NOTE =
;   OTHER INFORMATION: Synthetic Construct
US-10-237-302-7
```

```
  Query Match             51.9%;  Score 29.6;  DB 15;  Length 91;
  Best Local Similarity   59.4%;  Pred. No. 23;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db         18 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 49


RESULT 9
US-10-237-302-8/c
; Sequence 8, Application US/10237302
; Publication No. US20030119182A1
; GENERAL INFORMATION:
;  APPLICANT: AlphaVax, Inc.
;  APPLICANT:  Jonathan F. Smith
;  APPLICANT:  Kurt I. Kamrud
;  APPLICANT:  Jonathan O. Rayner
;  APPLICANT:  Sergey A. Dryga
;  APPLICANT:  Ian J. Caley
;  TITLE OF INVENTION: ALPHAVIRUS REPLICON VECTOR SYSTEMS
;  FILE REFERENCE: 01113.0002U2
;  CURRENT APPLICATION NUMBER: US/10/237,302
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 60/317,722
;  PRIOR FILING DATE: 2001-09-06
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence; NOTE =
;   OTHER INFORMATION: Synthetic Construct
US-10-237-302-8

  Query Match             51.9%;  Score 29.6;  DB 15;  Length 99;
  Best Local Similarity   59.4%;  Pred. No. 23;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db         78 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 47


RESULT 10
US-09-190-246-4
; Sequence 4, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT: Li, Xiaomao
;  APPLICANT: Klein, Michel
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: Parrington et al.
;  CURRENT APPLICATION NUMBER: US/09/190,246
;  CURRENT FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 8100
;   TYPE: DNA
;   ORGANISM: Semliki Forest virus
US-09-190-246-4

  Query Match             51.9%;  Score 29.6;  DB 10;  Length 8100;
  Best Local Similarity   59.4%;  Pred. No. 37;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db       8032 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 8063


RESULT 11
US-09-190-246-1
; Sequence 1, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT: Li, Xiaomao
;  APPLICANT: Klein, Michel
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: Parrington et al.
;  CURRENT APPLICATION NUMBER: US/09/190,246
;  CURRENT FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 15538
;   TYPE: DNA
;   ORGANISM: Semliki Forest virus
US-09-190-246-1

  Query Match             51.9%;  Score 29.6;  DB 10;  Length 15538;
  Best Local Similarity   59.4%;  Pred. No. 40;
  Matches   19;  Conservative   12;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 UCCACCUCCUCGCGGUNNNDNNUGGGCAUGCG 35
              :|||||:||:|||||::::::::|||||: ||
Db      12479 TCCACCTCCTCGCGGTCCGACCTGGGCATCCG 12510


RESULT 12
US-09-918-995-2716/c
; Sequence 2716, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2716
;   LENGTH: 466
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(466)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2716

  Query Match             51.2%;  Score 29.2;  DB 10;  Length 466;
  Best Local Similarity   50.0%;  Pred. No. 37;
  Matches   23;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGG 46
              |||: | ||:||:|| |  :::::::|| ||:|  ||: ||| :||
Db         99 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGG 54


RESULT 13
US-09-918-995-10300/c
; Sequence 10300, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
```

```
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10300
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(487)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10300

  Query Match             51.2%;  Score 29.2;  DB 10;  Length 487;
  Best Local Similarity   50.0%;  Pred. No. 37;
  Matches   23;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGG 46
              |||: | ||:||:|||  :::::::|| ||:|  ||: ||| :||
Db        196 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGG 151


RESULT 14
US-10-109-860-6/c
; Sequence 6, Application US/10109860
; Publication No. US20020142421A1
; GENERAL INFORMATION:
;  APPLICANT: SHAO, Wei et al
;  TITLE OF INVENTION: ISOLATED HUMAN AMINOACYLASE, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN AMINOACYLASE, AND USES THEREOF
;  FILE REFERENCE: CL001179DIV
;  CURRENT APPLICATION NUMBER: US/10/109,860
;  CURRENT FILING DATE:  2002-04-01
;  PRIOR APPLICATION NUMBER: 09/814,951
;  PRIOR FILING DATE: 2001-03-23
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-109-860-6

  Query Match             51.2%;  Score 29.2;  DB 14;  Length 601;
  Best Local Similarity   50.0%;  Pred. No. 38;
  Matches   23;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGG 46
              |||: | ||:||:|||  :::::::|| ||:|  ||: ||| :||
Db         77 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGG 32


RESULT 15
US-10-342-887-515/c
; Sequence 515, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 515
;   LENGTH: 1415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-515

  Query Match             51.2%;  Score 29.2;  DB 13;  Length 1415;
  Best Local Similarity   50.0%;  Pred. No. 42;
  Matches   23;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGUCCACCUCCUCGCGGUNNNDNNUGGGCAUGCGGCUUCGCAUGG 46
              |||: | ||:||:|||  :::::::|| ||:|  ||: ||| :||
Db         93 GGGTGCTCCTCCTCGGGACCCTTGCTGGTCATGGCGCTGCGCGTGG 48


Search completed: May 19, 2004, 04:59:42
Job time : 421 secs
```